How Many People Can China Feed?

Assessing the Impact of Land and Water Constraints

by

Amy Beth Watson

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Civil and Environmental Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2004

Author.....

Department of Civil and Environmental Engineering
May 17, 2004

Certified by.....

Dennis B. McLaughlin
H.M. King Bhumibol Professor of Water Resource Management
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by.....

Heidi Nepf
Chairman, Departmental Committee on Graduate Students

How Many People Can China Feed? Assessing the Impact of Land and Water Constraints

by
Amy Beth Watson

Submitted to the Department of Civil and Environmental Engineering
On May 17, 2004, in partial fulfillment of the
requirements for the degree of
Master of Science in Civil and Environmental Engineering

Abstract

Land and water resources are becoming increasingly scarce in China, threatening the nation's ability to feed its growing population. The limitations of these resources must be considered simultaneously to determine China's ability to produce food. In this thesis we present an optimization model to identify the maximum population that can be sustainably supported subject to land and water constraints. This model can be used to inform water resource management decisions.

The optimization model specifies the spatial allocation of cropland and water, subject to various physical constraints. As more land is used for crops, more water is consumed by agriculture. The model's water constraints are based on steady-state, annual water balances for major river basins and precipitation and evapotranspiration climatology. The movement of water is constrained by a coarse resolution stream network within each river basin. Food produced within each river basin may be consumed anywhere within China. The model's land constraints are based on reported values for total and irrigated cropland. The irrigated cropland constraint can be relaxed to examine production increases due to possible expansion of irrigated infrastructure beyond current levels.

The use of this model is demonstrated with preliminary data. The results suggest that China can support 693 million people sustainably with its resources, assuming current levels of crop imports. Expanding irrigation infrastructure to all cropland enables China to support 828 million people. The model proves to be highly sensitive to oil crop consumption and crop growing period inputs. Sensitivity to data inputs contributes to the uncertainty of model results. Further refinement of the model and improved data should result in improved population predictions.

Thesis Supervisor: Dennis B. McLaughlin
Title: H.M. King Bhumibol Professor of Water Resource Management
Professor of Civil and Environmental Engineering

Acknowledgments

First I'd like to thank my advisor, Dennis McLaughlin, for giving me the opportunity to work on such an exciting project and providing me with guidance through its entirety. The work would not have been possible without funding from the MIT/ETH/UT Alliance for Global Sustainability.

Many people contributed to the work that is presented in my thesis. Wes Dripps took the time to explain the project and hand the reins over to me, so to speak, when I began in the fall of 2002. Sal Scaturro provided personal insight to the problems in China and gave us a reality check on the amount of arable land in southern provinces. Shuang Hu worked with Liz Walker in the spring and summer of 2003 to convert the first version of the model pre-processor into a more comfortable coding language. Hui Li spent time developing the optimization model code during the summer of 2003. Suzanne Young spent the summer of 2003 developing the dietary demands discussed in section 4.4.2. Suzanne also worked with Shuang to translate data from Chinese texts. Piyatida Hoisungwan joined the project in the fall of 2003 and is largely responsible for the addition of the Degree Days method and the related data of section 4.4. Marine Herrmann began analyzing the preliminary results of the optimization model in January of 2004 and took responsibility for many of the changes to the model pre-processor as well as running the final trials of the optimization model presented in Chapter 5. The project could not possibly have progressed as far as it has without the help and contribution of all these people.

Special thanks are due to Enrique Vivoni for suggesting I read the Olivera and Raina (2003) article when I presented to him the challenge of water movement in our model. Without this chance encounter, our model would not have the hydrologic component that I feel is so crucial to its validity.

I want to thank everyone in Parsons Lab who has contributed to the successful completion of this work. I'd especially like to thank office mates, Marine Herrmann, Homero Flores-Cervantes, and David Rodriguez Gonzalez who did their best to keep me focused and forced me to take sanity breaks (even if I resisted).

I am very grateful to Chris Johns for all the support he has given over the past two years. His presence has made this a much more pleasant journey.

Last but not least I would like to my family for always showing their support and belief that I can do anything that I put my mind to. They have helped keep me grounded and focused on what is important in life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Water and land are arguably the two most important resources contributing to agricultural production. They are equally vital for the survival of life on this planet. Agriculture requires a significant amount of water; it is estimated that more than 85 percent of the world's fresh water resources are used for agriculture (Shiklomanov, 1993). Currently, about 10 percent of the earth's land surface is used as cropland, and 16 percent of this land is supplied with irrigated water (Postel, 1993). Irrigation allows water scarce land to be agriculturally productive land, filling the gap between natural water supplied through precipitation and the water demand of crops. However, in many parts of the world water shortages and arable land loss threaten food security. As global population increases, even relatively rich nations may feel the affects of these shortages.

In China, water scarcity and arable land loss pose serious threats to agricultural self-sufficiency, a goal of the nation's agriculture policy for many years (Huang, 1998). China is a relatively water-rich country, but the basic statistics on China's water resources do not tell a complete story. The nation receives over 2800 km^3 of renewable water resources every year and uses approximately 630 km^3 annually (Aquastat, 2000). Water scarcity is a problem because 80% of China's water resources are in the south and two-thirds of its cropland is in the north (Brown and Halweil, 1998). Furthermore, 83% of China's irrigated land occurs in the semi-arid region of the North China Plain (World Bank, 2001). Recent reports on the issue of water scarcity focus on the North China Plain (NCP), an area surrounding Beijing that is one of the most agriculturally productive regions in all of China. The NCP produces about half of the nation's wheat and one third of its maize (Kendy et al, 2003). However, it appears that this region has

been producing beyond its sustainable abilities. Withdrawals from groundwater sources have led to sharply falling water tables and land subsidence in the NCP. Brown and Halweil (1998) report that the water table has been falling by about 1.5 meters per year for quite some time. The World Bank (2001) estimates that the water tables in the three major river basins of the NCP have dropped 50 meters in some places. These water table declines were first noticed in the 1970's but agriculture flourished in spite of this evidence (Zhu and Zheng, 1983). Land subsidence occurs when water is removed from the subsurface at a rate higher than natural recharge, leaving voids in the soil matrix, which in time collapse causing the surface to subside. In addition to land subsidence, this overdrafting or "mining" of groundwater leads to salt water intrusion of aquifers, permanently contaminating limited fresh water resources. Water is such a heavily used resource in the NCP that streamflows have, periodically, ceased. Brown and Halweil (1998) suggest that if irrigated cropland were converted to rainfed cropland due to the limited availability of water, crop yields would decrease by 50-67%. It would, however, reduce the NCP's agricultural water usage to a sustainable point.

A growing population and ever changing water demands exacerbate the water shortage problem on the North China Plain. As the country becomes more industrialized a larger percentage of water will be used for industrial purposes, leaving less water available for agriculture. However, as population increases, so does food demand, which requires an increase in agricultural productivity. Currently China's water consumption for agricultural purposes is around 427 km^3 per year, but it is estimated to grow to 665 km^3 per year (Brown and Halweil, 1998). It is up to China's policy makers to determine the best way to manage this finite resource. According to Smil (1995), in the 1980's irrigation water was significantly under-priced at approximately 5-20% of the actual cost.

Arable land loss is another threat to China's agricultural capability. While it is difficult to obtain an accurate figure of China's cropland, many sources suggest that the land devoted to agriculture is on the order of 130 million hectares. Cropland is more

abundant in the North, but approximately 0.5 million hectares of cropland have been lost every year to other uses since the early 1980's (Smil, 1995). If this rate is maintained, 10-15 million hectares of arable land will be converted to other uses by the year 2020 (Smil, 1995). These losses are hard to control as they are typically the result of farmers converting their land for other horticultural uses or selling their land to live a seemingly more profitable urban life. The government has tried to impose regulations on the conversion of cropland to other uses, but so far it appears to have had little affect.

Another serious factor contributing to arable land loss is desertification. According to the World Bank (2001), there are 331 million hectares of land in China that are prone to desertification and 262 million hectares where it is actually occurring. It is difficult to estimate how much of this land is cropland, but with farm land abandonment on the rise due to insufficient water supply, it is not hard to imagine that some of this land is, or once was, cropland. Cropland is also lost due to desertification because some farmers are paid to plant trees on their land to prevent the spread of the growing deserts (Brown, 2004).

Despite losses in arable land and an uneven distribution of water resources, China's population has continued to grow by approximately 1% annually. The United Nations (UN) estimates that by 2025, China's population will exceed 1.5 billion people. As of the beginning of 2004, China's population is about 1.3 billion people. Feeding an additional 200+ million people with shrinking resources is a challenge that many at the World Watch Institute believe China cannot meet on its own. According to Smil (1995), if China imports 100 million tons of grain each year, the equivalent of half the global annual grain shipments, grain prices will double. And if they must import 200 million tons of grain it is unlikely that any increase in grain prices will enable them to meet this demand. It is, therefore, of utmost importance that the Chinese government decides how best to meet the dietary demands of their growing population.

It is an irrefutable fact that water scarcity and arable land loss exist in China. The question is how much of an affect will limited land and water resources have on agricultural production. Smil (1995) concludes that "better management, better pricing,

better inputs and better environmental protection can lead to enough additional food from China's agroecosystems to provide decent nutrition during the next generation." Others, such as Brown (1995), believe that China is already producing beyond what is considered sustainable production and will have to depend on imports from other nation's to make up the difference between domestic production and demand. If this desperate conclusion is true, global food prices will increase and poorer nations in Africa and Asia will suffer the greatest consequences, as less people will be able to afford the ever more costly food supply.

This research was conducted to examine the controversy over China's ability to be self-sufficient in food production. Many scientists and economists have tried to address this problem, but few (if any) have taken a thorough look at the factors influencing agricultural production in China. One study conducted by IIASA (Heilig, 1999) looked closely at changes in arable land, potentials for multi-cropping, urbanization and changes in diet, but essentially ignored the limitations of water on food production. Other publications, such as Smil (1995), report only national figures, which don't take into account the spatial distribution of China's resources and how the distribution might affect productivity. As is often the case, it is the short-comings of previous work that motivated this research.

The objective of this research is to explore the physical limitations of sustainable agricultural production in China. To that end, this paper will discuss an optimization model designed to maximize the population fed by Chinese resources, including current import levels, subject to natural constraints on land and water. The model must account for agricultural production as well as consumption levels. It must also be designed with the spatial and temporal resolution needed to properly describe crop growth and resource availability.

Model results are only as good as the quality of the data input to them. Therefore, an additional component of model development is to compile a comprehensive dataset that quantifies all the variables in question. Obtaining adequate, valid data is a non-

trivial task, and is not taken lightly in this research. We attempt to determine if the data used in the model is adequate and if the data is of high quality and validity.

Finally, this research hopes to produce a preliminary estimate of the maximum population that can be supported by Chinese agricultural production. We also provide an estimate of the potential population supported if irrigation infrastructure were expanded throughout the country so that all arable land is irrigable.

In summary, the primary objectives of this research are to:

1) **Develop an optimization model to identify the maximum number of people fed with Chinese resources.**
2) **Construct a comprehensive data set describing land and warer resources in China.**
3) **Use this data set to obtain a preliminary estimate of the population capable of being supported by Chinese resources with current and potential levels of irrigation infrastructure.**

Chapter 2 provides a discussion of the fundamental principles behind the development of the optimization model used to simulate sustainable agriculture in China. A detailed discussion of all the model components is given in Chapter 3. This includes descriptions of how each model variable is calculated and the formulation of the objective function and constraining equations of the optimization. Chapter 3 also explains some of the assumptions and decisions made while constructing the model. Chapter 4 describes the global data used to describe the land and water resources in China as well as other agronomic parameters. In Chapter 5 we present the results of optimization model along with a comparison to statistical data on agricultural production. Chapter 6 is a discussion of the conclusions reached with this research and suggestions for future improvements of the model.

Chapter 2

Model Conceptualization

The main objective of this research is to answer the fundamental question posed in Chapter 1, how many people can China feed? To accomplish this, we are posing the question as an optimization problem in which the objective is to maximize the number of people fed subject to the physical limitations on the primary agriculture resources, land and water, in China. In this Chapter we explore the concepts and assumptions used to formulate the optimization problem. Chapters 3 and 4 will elaborate further on the details discussed in this Chapter as well as the data used in the model.

Inherent to this problem is the need to model agricultural production in China. As described in Chapter 1, production requires large amounts of fresh water as well as cropland. So the population supported by agriculture is actually dependent on the land dedicated to crop production and the water available to cropland. In order to determine how much food China can produce, and therefore how many people the nation can feed, we must determine how much water and land are available for agriculture. This can be done using the mathematical concepts summarized below in Figure 2-1.

$$\begin{gathered} Max\ People\ Fed \\ P - E(cropland, non-arable land) = R \geq 0 \\ cropland \leq arable\ land \\ People\ Fed = f(cropland) \end{gathered}$$

Figure 2-1: Summary of model concepts

Line 1 of Figure 2-1 represents the objective of the model, to maximize the population fed with Chinese resources. Line 2 is a basic water balance, which is used to determine the availability of water. In this simplistic form, total runoff, R, is calculated as the difference between precipitation, P, and evapotranspiration, E. Evapotranspiration is a function of the land use, primarily cropland and non-arable land, and serves as the connection between crops, land, and water. It increases as the amount of cropland increases, but cannot exceed precipitation. Line 3 specifies the constraint that only arable land can be used for cropland. Line 4 represents the idea that, assuming adequate conditions for crop growth are met, the number of people fed is a direct function of the amount of land used to grow crops. In this formulation either water (Line 2) or land (Line 3) limits the number of people fed.

2.1 Water Constraints

Water constraints are a consequence of the water balance, which is an accounting of water mass over some control volume. It consists of defining a control volume and determining the inputs and outputs of the control volume. For the purposes of this research we wanted to keep the water balance as simple as possible while resolving important spatial and temporal variations.

In particular, the water balance must be able to describe seasonal and regional variations in crop production. Fortunately, the availability and quality of hydrologic data has been increasing, in large part, to advancements in remote sensing technology. However, it is important to remember that fine resolution modeling requires a significant amount of data (especially for a large study region such as all of China), which correlates to a significant amount of computation. Since the program used to write the optimization model has computational limitations, the goal of the model conceptualization is to determine adequate spatial and temporal scales that will minimize computation but still capture the affects of variables that change at small scales.

2.1.1 Spatial Scale

The spatial scale of the model defines the control volume over which the water balance is calculated. It is obvious that modeling a water balance at the country-scale is too coarse as it does not take into account the uneven distribution of water resources described in Chapter 1. River basins present a logical control volume for the model. We have ignored the affects that groundwater basins have on surface basin boundaries by assuming that there is no groundwater flow between basins. Since there are no lateral inflows from one basin to another, basins can be considered self-contained and are nature's control volume for water. Moreover, there are about 20 sizable river basins in China, a manageable number of units to model.

In order to construct an accurate water balance at the river basin scale we need to have accurate estimates of basin total precipitation and evapotranspiration. Since precipitation and evapotranspiration vary at scales much smaller than river basins it is necessary to subdivide basins into smaller units. The smallest unit of the model is defined as a 0.5^{o} x 0.5^{o} unit, called a pixel. The water balance must be performed for each pixel and each basin within China. Basins are delineated at 0.5^{o} x 0.5^{o} resolution, as shown in Figure 2-2. The technique for deriving these basin boundaries is discussed in Section 3.3.2 of the following Chapter.

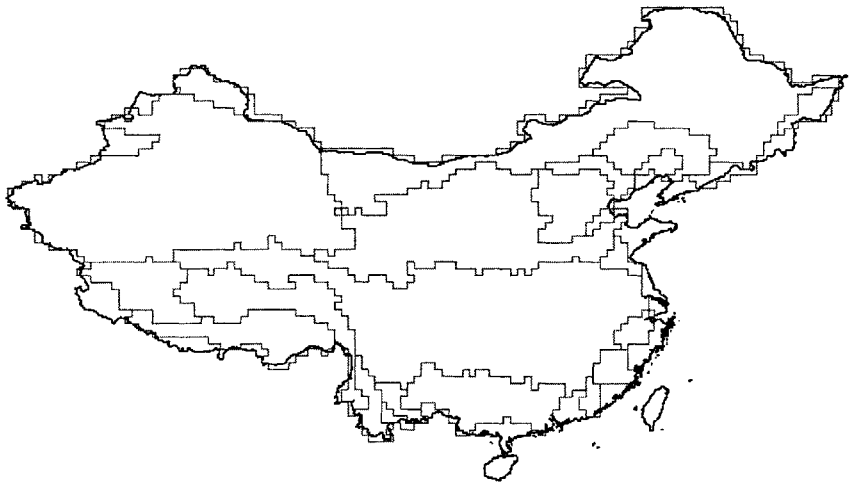


Figure 2-2: River basins of China at 0.5^{o} x 0.5^{o} resolution.

2.1.2 Temporal Scale

Climate and hydrology exhibit annual periodicity. Due to this periodicity the climate repeats itself each year in seasons. Farmers have used this knowledge for thousands of years to make agronomic decisions. A climatology averages climate over many years, allowing us to make an estimate of the climate in a typical year and also to assume that China's water resources are in a steady-state over the long term. Steady-state is essential for sustainable water use since it implies that changes in storage, such as aquifers and reservoirs, are negligible. We use climatology information to apply an annual time scale to the water balance and crop production. This resolution helps minimize computation in the optimization model. However, the seasonal variation of water supply and crop growth cannot be neglected. Therefore, the model assumes that water and food can be stored intra-annually. Intra-annual storage of food, especially grains, is not difficult as it is part of current practice in most parts of the world. Intra-annual storage of water is more difficult as it requires local dams or reservoirs at 0.5^{o} x 0.5^{o} spatial resolution (approximately 50km by 50km). Since China is a heavily irrigated country we assume that this infrastructure already exists.

In order to model crop production, the model must include a smaller time resolution to account for the seasonality of crop growth. Crops are grown in seasons, so the water needed to grow crops changes throughout the year. Moreover, precipitation and evapotranspiration rates exhibit seasonal variations. The water balance must be able to capture the changes in water fluxes at a seasonal, if not monthly, time scale.

2.2 Land Constraints

Precipitation and evapotranspiration vary over time and space, but as shown in Figure 2-1 evapotranspiration is also a function of the land cover. Vegetated land evapotranspires at different rates than bare soil or pavement, for instance.

This is because evapotranspiration is the combination of two processes, evaporation and transpiration. Evaporation is a measure of the atmosphere's ability to hold water. Transpiration is a process by which plants release water to the atmosphere. Both processes are affected by climatic conditions, but transpiration is also dependent on the type of plant. As a result, the model must take into account the variability between different land use evapotranspiration rates.

2.2.1 Evapotranspiration by Land Use

There are many classifications of land use, but since our model is concerned with modeling agricultural production we will consider two basic types of land use; cropland and non-arable land. Non-arable land is assumed to have uniform properties across the country, so evapotranspiration is only dependent on time and location of the land. Cropland, however, should be further defined by the type of crops grown on the land and the timing of their growth. This is done by defining crop sequences in which one, two, or three crops are grown in one calendar year. Six crops are considered in the model because they make up a significant portion (90%) of the Chinese diet, but a seventh crop is included due to its unique sensitivity to temperature. These seven crops can be combined to create 301 crop sequences: 7 single-crop, 42 double-crop, and 252 triple-crop sequences (see section 4.4.4 for a discussion of crop sequences). The model determines when each crop in a crop sequence is grown and calculates the evapotranspiration rate for each crop season. Summation of the seasonal evapotranspiration rates results in annual evapotranspiration rates. Annual rates allow water allocation to be done on an annual time scale. This method of calculating evapotranspiration allows for the seasonality of water use to be accounted for while still maintaining the larger annual temporal resolution.

2.2.2 Water Supply by Land Use

Irrigated vs. Dryland

Agriculture production is limited by the amount of land dedicated to crop growth as well as the amount of cropland equipped for irrigation. Approximately 47% of China's cropland is irrigated (Postel, 1993). Crops that are not grown on irrigated land must obtain enough water to sustain growth from precipitation alone. This is known as dryland agriculture. Using the climatology of a typical year, the model will not plant a crop under dryland conditions unless that land typically receives enough water for the crop. Irrigated land can grow crops so long as the water needed to grow the crop can be provided throughout growth. The model will not grow a crop on irrigated land unless, on average, the water supply is adequate for that crop.

Arable land loss makes the possibility of cropland area expansion unlikely. Furthermore, it is difficult to quantify or predict the amount of arable land that will be lost in the future. Therefore, current cropland data is used as an upper bound on the amount of arable land available now and in the future. We use irrigated land data to make an estimate of the population China can currently feed, but future expansion of irrigation could increase the nation's productivity. This possibility is explored in Chapter 5.

Non-arable vs. Cropland

Non-arable land is handled differently from cropland in the model. Vegetation that grows on non-arable land requires water to sustain growth, just like food crops. However, the model is not concerned with sustaining growth of vegetation on non-arable land. Vegetation on non-arable land is not managed, so the plants that grow on it must be sustained with the local climatology, similar to dryland agriculture. If precipitation falls short in a year, the non-arable land vegetation could die. The model ensures that cropland vegetation will not die

due to water shortages because crops are not planted if the water supply is inadequate.

A summary of the spatial- and temporal-scales used in the model are provided in Tables 2-1 and 2-2. The use of each scale becomes more obvious in the following Chapter.

Scale	Purpose
China	food suuply allocation
Basin	water fluxes
Pixel	water fluxes, land use
Provincial	comparison of crop production

Table 2.1: Spatial-scales used within the optimization model.

Scale	Purpose
Annual	water supply allocation
Monthly	seasonal variations in crop grouwth
Daily	interpolated from monthly to define crop growing period

Table 2.2: Temporal-scales used within the optimization model.

With the general framework of the model defined we are ready to explore the optimization model in detail. Chapter 3 will present a detailed description of the optimization model and all of its components. Chapter 4 will then discuss the data used to quantify all the variables and inputs of Chapter 3.

Chapter 3

Detailed Model Description

The general framework of the optimization model was discussed in Chapter 2 so this chapter will focus on a detailed discussion of the model components. This chapter will simultaneously describe the steps necessary for making a complete run of the model.

3.1 Flow Chart

The procedure for running the optimization model is diagrammed in the flow chart of Figure 3-1, below.

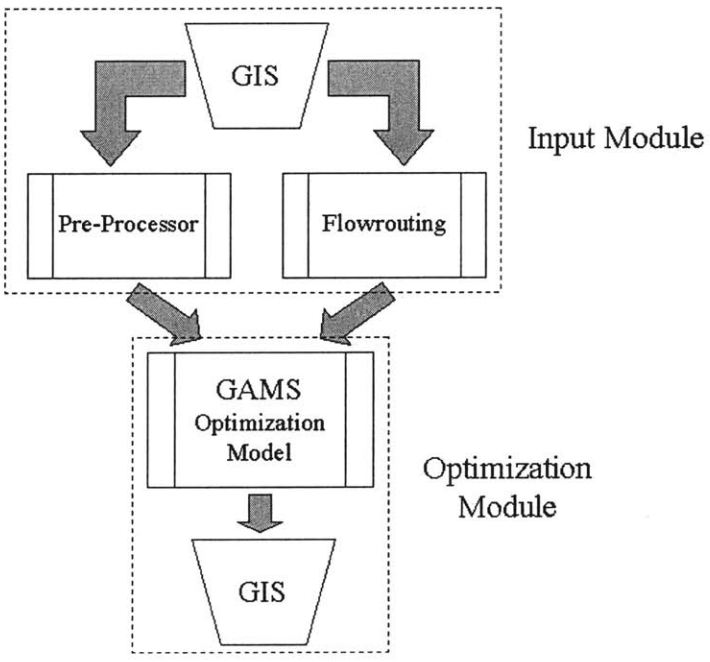


Figure 3-1: Flow chart of agricultural optimization model.

The first portion of the optimization model is the input module. Raw data is obtained and processed in GIS for input to the Pre-Processor MATLAB code and

Flowrouting program. The Pre-Processor calculates crop sequence annual evapotranspiration rates and writes formatted input data files for the optimization module. The Flowrouting program delineates 0.5° x 0.5° resolution river basins. Information from the input module is sent to the optimization module. In the optimization module, the GAMS model is used to obtain the maximum population supported by land and water resources, and outputs files describing crop production and water use in each pixel. These files are then processed in GIS to obtain production estimates at the pixel- provincial-, and national-scale.

3.2 Optimization Module

The optimization module contains the GAMS optimization model as well as post-processing of the results and visualization in GIS. In this section both components of the optimization module are discussed, beginning with a detailed discussion of the GAMS optimization model formulation.

3.2.1 GAMS

The optimization model is written in the General Algebraic Modeling System (GAMS). GAMS is an ideal environment for modeling linear, non-linear, and mixed integer optimization problems as it can handle large, complex systems (GAMS, 2004). The optimization model described here proves to be a very large, linear problem. Optimization problems consist of an objective function subject to any number of constraints with decision variables serving as the link between the objective and constraints. In the following sections the objective, decision variables, constraints, and other data inputs are presented. A complete version of the optimization code is provided in Appendix A.

Objective Function

Maximize: **N**, the number of people supported by agricultural production in China.

Decision Variables

Decision variables are unknowns of the problem and are determined by the model when solving for the objective function. The decision variables for this model are:

$LI_{i,s}$: irrigated cropland devoted to crop sequence s in pixel i [10^6 ha]

$LD_{i,s}$: dryland cropland devoted to crop sequence s in pixel i [10^6 ha]

R_b: runoff in basin b [10^6 m^3]

E_i: actual evapotranspiration from pixel i [10^6 m^3]

W_b: total municipal and industrial consumption of water in basin b [10^6 m^3]

$M_{i,c}$: total production of crop c in pixel i [10^9 kg]

N: number of people supported by agricultural production in China

The primary decision variables of the model are $LI_{i,s}$ and $LD_{i,s}$, land devoted to agriculture. The model decides how much land is used for individual crop sequences in each pixel under dryland or irrigated conditions. The land decision variables appear in several constraining equations.

Secondary decision variables, such as the hydrologic variables basin runoff, evapotranspiration, and municipal/industrial water consumption, appear in the water balance constraint equations. The production variable appears in the food balance constraint.

Constraints

Constraints are equalities or inequalities that restrict the values of the decision variables. Inputs, which also appear in constraints, will be defined in the next section.

Water Balance Constraints

Annual Water Balance for River Basin *b*:

$$P_b - E_b - W_b - R_b = \frac{dS_b}{dt} \tag{3.1}$$

where P_b is the total basin precipitation, E_b is the total basin actual evapotranspiration, W_b is the municipal and industrial water consumption of the basin, R_b is the runoff out of the basin, and S_b is the storage of water in basin b. As mentioned in Chapter 2, water use is sustainable if the water balance is in steady-state. At steady-state, the change in storage is negligible over time-scales greater than one year, so we assume $\frac{dS_b}{dt}$ is zero and Equation 3.1 becomes,

$$P_b - E_b - W_b = R_b \tag{3.2}$$

Total basin precipitation is calculated as,

$$P_b = \sum_{i=1}^{B_b} (P_i)(A_i) \tag{3.3}$$

where P_i is the precipitation in pixel i, A_i is the total area of pixel i, and B_b is the total number of pixels within basin b. A similar equation gives the total basin evapotranspiration,

$$E_b = \sum_{i=1}^{B_b} E_i \tag{3.4}$$

where E_i is the actual evapotranspiration (AET) from pixel i. AET is different for different land uses (discussed in Section 3.3.3), so E_i is expanded to:

$$E_i = E_{i,a} + E_{i,na} \tag{3.5}$$

where $E_{i,a}$ is AET from arable land, $E_{i,na}$ is AET from non-arable land, subscript a refers to arable land and subscript na refers to non-arable land. If non-arable is included as a crop sequence, this can be written as,

$$E_i = \sum_{s=1}^{302} [(E\_I_{i,s})(LI_{i,s}) + (E\_D_{i,s})(LD_{i,s})] \tag{3.6}$$

where $E\_I_{i,s}$ and $E\_D_{i,s}$ are the actual evapotranspiration rates of crop sequence s in pixel i for irrigated land and dryland, respectively. Equation 3.5 sums all the evapotranspiration from all 302 possible sequences (301 crop sequences plus non-

arable) for both irrigated and dryland agriculture. $E_{i,a}$ and $E_{i,na}$ can also be considered separately,

$$E_{i,a} = \sum_{s=1}^{301} \left[\left(E\_I_{i,s} \right) \left(LI_{i,s} \right) + \left(E\_D_{i,s} \right) \left(LD_{i,s} \right) \right] \tag{3.7}$$

$$E_{i,na} = \left(E\_I_{i,na} \right) \left(LI_{i,na} \right) + \left(E\_D_{i,na} \right) \left(LD_{i,na} \right) \tag{3.8}$$

where for $E_{i,a}$, AET is summed over only the 301 crop sequences alone and for $E_{i,na}$, AET is summed only for non-arable land.

Municipal and industrial water use, W_b, is calculated according to the following,

$$W_b = \left(N_b \right) \left(W \right) \tag{3.9}$$

where N_b is the population in basin b and W is the per capita municipal and water consumption factor. N_b is based on the national population, N, and a basin population distribution factor, F_b, as follows,

$$N_b = \left(N \right) \left(F_b \right) \tag{3.10}$$

Substituting Equations 3.2-3.9 into Equation 3.1, the basin water balance becomes:

$$\sum_{i=1}^{B_b} \left[P_i - \sum_{s=1}^{302} \left(ET\_I_{i,s} \right) \left(LI_{i,s} \right) + \left(ET\_D_{i,s} \right) \left(LD_{i,s} \right) \right] - \left(N \right) \left(W \right) \left(F_b \right) = R_b \tag{3.11}$$

Annual Water Balance in Pixel i:

Assuming steady-state, the change in storage in a pixel is negligible and the water balance in pixel i is,

$$P_i + R_{i,entering} - E_i - R_{i,exiting} = 0 \tag{3.12}$$

$R_{i,entering}$ is the sum of all the runoff generated from tributary pixels entering pixel i,

$$R_{i,entering} = \sum_{n(i)=1}^{T_i} R_{i,n(i)} \tag{3.13}$$

where $n(i)$ refers to the n^{th} pixel tributary to pixel i, T_i is the total number of tributary pixels to pixel i, and $R_{i,n}$ is the runoff generated in tributary pixel $n(i)$. $R_{i,exiting}$ is the

remainder of water that is not used in pixel i for evapotranspiration. This runoff flows into the immediate downstream pixel either on the surface (through streamflow) or the subsurface (through groundwater flow).

Substituting Equation 3.5 and 3.12 into Equation 3.11 results in the complete form of the pixel water balance,

$$P_i + \sum_{n=1}^{T_i}(R_n) - \sum_{s=1}^{302}\left[\left(E\_I_{i,s}\right)\left(LI_{i,s}\right) + \left(E\_D_{i,s}\right)\left(LD_{i,s}\right)\right] - R_{i,exiting} = 0 \qquad (3.14)$$

Land Balance Constraints

Total Area for Pixel *i*: $$\sum_{s=1}^{302} LI_{i,s} + LD_{i,s} = A_i \qquad (3.15)$$

where the summation is over all crop sequences, including non-arable land.

Arable Land: $$\sum_{s=1}^{301}\left(LI_{i,s} + LD_{i,s}\right) \le A_{i,crop} \qquad (3.16)$$

Irrigable Land: $$\sum_{s=1}^{301} LI_{i,s} \le A_{i,irr} \qquad (3.17)$$

where A_i is the total land in pixel i, $A_{i,crop}$ is available cropland, $A_{i,irr}$ is available irrigated land, and the sequences summed in (3.16) and (3.17) are only the crop sequences.

Production-Consumption Constraints

Consumption over all pixels:

$$\sum_{i=1}^{I}\left(M_{i,c}\right) + S_c = (N)(C_c) \qquad (3.18)$$

where $M_{i,c}$ is the production of crop c in pixel i, I is the total number of pixels in China, S_c is the average annual net import of crop c into China (at current levels), and C_c is the per capita consumption of crop c in the typical Chinese diet and N is the total population in China. $M_{i,c}$ is further expanded to:

$$M_{i,c} = \sum_{s=1}^{301} [(LI_{i,s} + LD_{i,s})(Y_{c,s})] \tag{3.19}$$

where $Y_{c,s}$ is the expected yield of crop c from sequence s. Equation 3.18 is then re-written as follows,

$$\sum_{i=1}^{I} \left(\sum_{s=1}^{301} [(LI_{i,s} + LD_{i,s})(Y_{c,s})] \right) + S_c = (N)(C_c) \tag{3.20}$$

The expression for total supply of crop c is on the left hand side of Equation 3.19 and the expression for total demand of crop c is on the right hand side of the equation. The two sides of the expression are equivalent to prevent excessive crop production.

Inputs

Input data are used to force the optimization model. The inputs of the optimization model follow the same subscript notation as the decision variables:

P_i: Precipitation [mm/yr]

$E\_I_{i,s}$: irrigated AET in pixel i for sequence s [mm/yr]

$E\_D_{i,s}$: dryland AET in pixel i for sequence s [mm/yr]

NOTE: For simplicity, non-arable is defined as a sequence so that a third AET variable is not needed. The calculation of AET for non-arable land is described in Section 3.3.3.

A_i: area of pixel i within Chinese border [ha]

$A_{i,crop}$: current cultivated land (cropland) in pixel i [ha]

$A_{i,irr}$: land equipped for irrigation, irrigated land, in pixel i [ha]

$Y_{c,s}$: yield for crop c in sequence s [kg/ha]

S_c: annual net import of crop c [10^9 kg/yr]

C_c: per capita crop consumption of crop c [kg/person/yr]

W_b: municipal and industrial water consumption in basin b [m^3/yr]

W: per capita municipal and industrial water consumption [m^3/person yr^{-1}]

F_b: basin specific population distribution [-]

Solving the Optimization Model

Many solvers can be used with GAMS, but CPLEX was chosen for this model as it is specifically designed to handle large, linear problems such as the one presented here. All equations in the model are solved simultaneously to define the feasible region of the problem. Each constraint above must be met to obtain a feasible solution.

3.2.2 Post-Processing

In order to study the affects of the physical limitations on agricultural production in China, it is necessary for the model to output more information than just the supported population (N). Other results that are interesting and informative are the basin runoff values (R_b), crop production at the pixel- ($M_{i,c}$), provincial- and national-scale, and evapotranspiration from arable land ($E_{i,a}$). These results can be compared with actual data specific to China.

Pixel-scale model results (e.g. $M_{i,c}$, $E_{i,a}$, and $E_{i,na}$) are output for each pixel in a text file so that they can be read by the Post-Processor, a program written in MATLAB. The Post-Processor takes the GAMS results for each pixel and writes them to an ASCII file, including header information, so that they are compatible with GIS. The results are converted to ESRI grid format using the ArcInfo ASCIIGRID command. Evapotranspiration results (E_i) can be compared to actual evapotranspiration data at the pixel-scale. Production results are compared with provincial statistical data, so pixel-scale results must be aggregated up to the provincial-scale. This is done with an area-weighted technique so that production results from pixels that are part of multiple provinces are divided according to their total area within each province. The comparison of model results to statistical data is conducted in Chapter 5.

3.3 Input Module

The first portion of the optimization model, the input module, performs many calculations and creates formatted input files for the GAMS model. The three main

components of the input module are GIS, the Pre-Processor MATLAB program, and the Flowrouting program. Each of these components are discussed in the following sections.

3.3.1 GIS

Geographic Information Systems (GIS) is a powerful tool for analyzing and manipulating spatially distributed data. GIS allows users to take raw data and project it onto a map surface, superimpose it with other data and perform many calculations. Since datasets come in varying map projections, extents, and units, GIS is needed in order to make comparisons between different formats of spatially distributed data. The following section describes how GIS is utilized in the input module.

Defining Study Area Extent

Data for three of the model inputs (precipitation, potential evapotranspiration, and temperature) are available globally in the 0.5° x 0.5° spatial resolution of the model pixels. However, the model only needs data that falls within the boundary of China. The cropland data set used in the model and discussed in the following Chapter, only covers the region from 73.5°E-134.5°E, 18.5°N-53.5°N, which includes most of China (missing data for Taiwain and southern tip of Hainan Island). Therefore the cropland dataset limits the total area in China that can be studied. All other gridded datasets are clipped to the same extent as the cropland dataset.

Gridded datasets, or grid files, are rectangles made up of rows and columns. The area studied in China is 70 rows by 122 columns, but not all the pixels in the grid are actually contained within the boundary of China. To determine which pixels are wholly or partially contained within China's border, a polygon coverage file with the same spatial extent and cell size of the gridded datasets is created in ArcInfo. This "gridded" polygon coverage is clipped with a China boundary coverage, described in Section 4.2.1, to determine which pixels on the grid are included in China. After clipping, any grid pixels with area greater than zero are deemed to be within the

border. The resulting coverage is referred to as the China boundary mask coverage, or simply the mask coverage, and is shown in Figure 3-2.

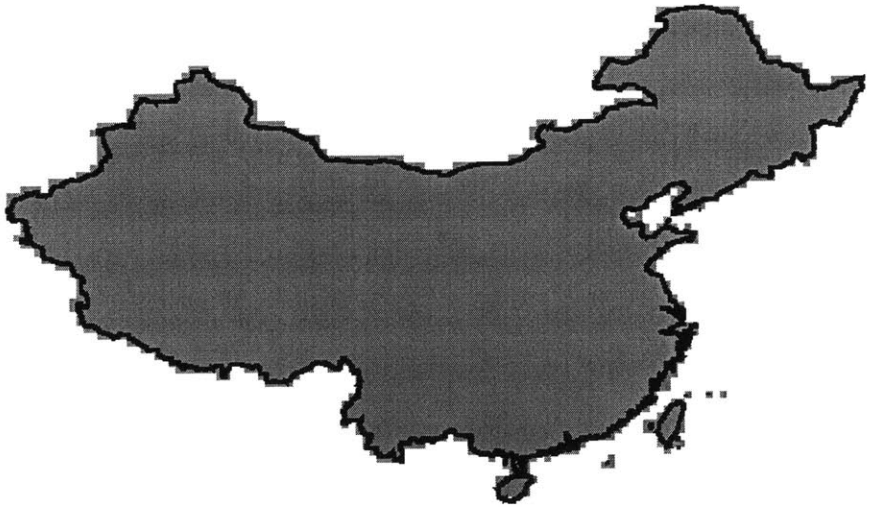


Figure 3-2: China mask coverage in gray, overlayed with China border.

GIS enables users to work in hundreds of different map projections. Each projection is useful for a specific purpose, which is the reason so many projections are needed. The data grid files used in this research were provided in geographic coordinates. The problem with working in geographic coordinates is that the area of each 0.5^{o} x 0.5^{o} grid pixel changes as you move in latitude. For the water balance the model needs to know the area of each grid pixel. In order to determine the total area of each pixel, the grid must be projected to a map projection that preserves area. The Lambert Azimuthal Equal Area projection was used for this purpose as it can be optimized for accuracy in China. We have determined that it is best to work in the original map projection of data because errors occur when data is projected to different coordinate systems. The extent of pixels in a different map projection do not always have the same extent as pixels in the original projection because the earth's spherical surface causes distortions. As a result, projected grid files data cannot be overlayed in precisely the same location as the original grid. To avoid the error inherent to projecting grid files, the China boundary coverage and the "gridded" mask coverage (not the grid file) are projected to Lambert Azimuthal Equal Area projection. Then the boundary coverage is used to clip the

gridded coverage. The total area of each pixel contained in China can be obtained from the clipped coverage (A_i - see Equations 3.2 and 3.14). This is especially important for boundary pixels that are only partially contained in China, but it is also needed to account for the latitudinal affects.

The mask coverage, which defines which pixels are contained within China and which pixels are not, is used to filter data grid files to remove any pixels that are not in China. The mask coverage is first converted to a grid file, with the value 1 representing pixels in China and 0 representing pixels outside of China. By multiplying the mask grid with the data grid files, all data outside of China is given the value 0 and all pixels within China retain their original value. The mask grid is used in the Flowrouting program, described in Section 3.3.2, and the Pre-Processor, which will be discussed in Section 3.3.3.

Writing Input Files for the Pre-Processor

GIS is used to write ASCII text files for input to the Pre-Processor. ASCII files consist of a 6-line header, which contains information on the spatial extent and value for no data pixels, and lines of space-delimited data. All gridded data, including monthly average precipitation, potential evapotranspiration, and temperature, as well as cropland and irrigated area, are projected to geographic coordinates, clipped to the same spatial extent, filtered to remove data outside China, and output as ASCII files by GIS.

3.3.2 Flowrouting

Flowrouting information is necessary for the model's water balance at both the basin- and pixel-scale. The model needs to know the tributary pixels to each pixel in the grid as well as all the pixels included in each 0.5^o x 0.5^o river basin. Traditionally, hydrologists use fine-resolution topographic data to delineate the boundaries of river basins. However, since 0.5^o x 0.5^o pixels are the smallest units of the model it is necessary to delineate river basins at this coarse resolution. This is a non-trivial issue.

Topographic data cannot be aggregated to the 0.5° x 0.5° resolution to obtain coarse resolution, or upscaled, river basins as this approach would not take into account the small-scale movement of the stream network. Determining upscaled river networks is a difficult task, but the demand for these networks is increasing as hydrologists and climatologists develop global-scale models.

The Network Tracing Method (NTM) developed by Olivera and Raina (2003) is the method employed to delineate gridded 0.5° x 0.5° resolution river networks for the optimization model. The shortcomings of previously developed methods include unbalanced flow direction distributions and lack of incorporating flow distance information into the upscaled networks (Olivera and Raina, 2003). The NTM was developed to address these limitations and create an overall improved method for delineating coarse resolution river networks.

Description of Network Tracing Method

The Network Tracing Method (NTM) uses fine-resolution topographic data to delineate gridded, coarse resolution river networks. The method enables the user to create a balanced flow distribution by varying a network defining parameter. NTM calculates and uses flow distance information to determine the direction of flow. This information helps distinguish meandering stream networks and is important for projects computing flow times.

The first step of the NTM is to derive a vector stream network with GIS. Topographic data in the form of a digital elevation model (DEM) is used to create a flow direction grid. For this research the Hydro1k, 1-km resolution DEM was used (see Section 4.1.4). The DEM is used to generate a flow direction grid. Flow direction grids define the direction of flow water would take from one fine-resolution pixel to another. Each pixel has eight neighboring pixels and water can flow to any of these. Arc determines the direction of steepest slope, and therefore most probable direction of flow, and creates a grid specifying the direction of flow for each pixel. The flow

direction grid is created at the ArcINFO grid prompt with the FLOWDIRECTION command.

The flow direction grid is then used to create a flow accumulation grid. This is also done at the ArcINFO grid prompt, using the FLOWACCUMULATION command. Arc analyzes the flow direction grid and counts the total number of pixels flowing to each pixel in the fine-resolution grid. Along streamlines and river reaches the flow accumulation is large, so these areas have high values. River outlets have the highest values in the flow accumulation grid because all the upstream pixels are flowing to those points. Figure 3-3 shows the flow accumulation grid.

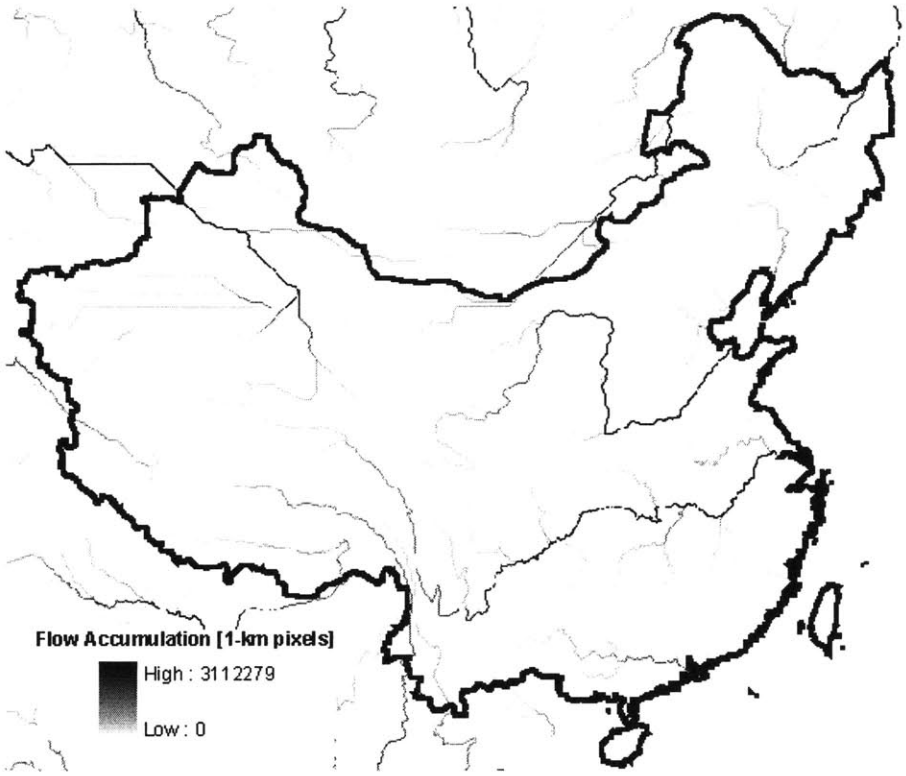


Figure 3-3: The flow accumulation grid makes major rivers clearly visible within China.

The flow direction and accumulation grids are used to construct a vector stream network. An Arc Macro Language script (AML) is used at the ArcINFO prompt to automate the derivation of the fine-resolution stream network. In order to run the AML, the user must specify the regional extent of the flow accumulation and direction

grids (in Lambert Azimuthal Equal Area projection) as well as the stream network threshold.

The stream network threshold defines how many fine-resolution grid pixels must flow to a pixel in order for it to be included in the vector stream network. Determining the threshold value is not a trivial task; a value that is too high will result in an incomplete stream network with some 0.5° x 0.5° pixels lacking a stream network, and a value that is too low will result in an overly complex stream network, which will greatly increase the time needed to run the NTM Visual Basic (VB) executable. Olivera recommends starting with a value of 1000 and working iteratively downward (personal communication, June 20, 2003). The user should choose the highest threshold value possible that allows the stream network to extend into all the necessary coarse resolution grid pixels (in our case, pixels within China).

Figure 3-4 shows a stream network developed by the AML with a stream network threshold value of 1000. A much more detailed stream network, developed with a threshold value of 10, was used to delineate the upscaled river basins of this model, but this simplified version is easier to interpret, visually.

Figure 3-4: Vector stream network of eastern China derived with a stream threshold value of 1000.

The next step of the AML is to intersect the fine-resolution stream network with the coarse resolution grid (in our case, 0.5° x 0.5°). The endpoint of each section of the intersected network is called a node, and nodes on the boundary of grid pixels are exit nodes. Finally, the AML writes a formatted text file containing the stream network information for input to the NTM VB executable. The NTM VB executable uses the stream network text file to calculate the distance from one exit node to the next. This distance is referred to as the stream reach length. If the stream network flows out of pixel "A" into pixel "B" and the reach length within pixel "B" is greater than a specified length, then pixel "B" is deemed the downstream pixel of pixel "A". The specified length is called the upscaled flow direction threshold and is input by the user. The upscaled flow direction threshold is directly related to the side-to-corner ratio. This ratio specifies the number of grid pixels where water flows through the sides and through the corners to the downstream pixel. A ratio value of 59:41 corresponds to a balanced flow distribution, according to the findings of Olivera et al. (2002). The user can obtain the ideal ratio by varying the upscaled flow direction threshold value. Figure 3-5, taken from Olivera and Raina (2003) demonstrates the basic principles of the NTM.

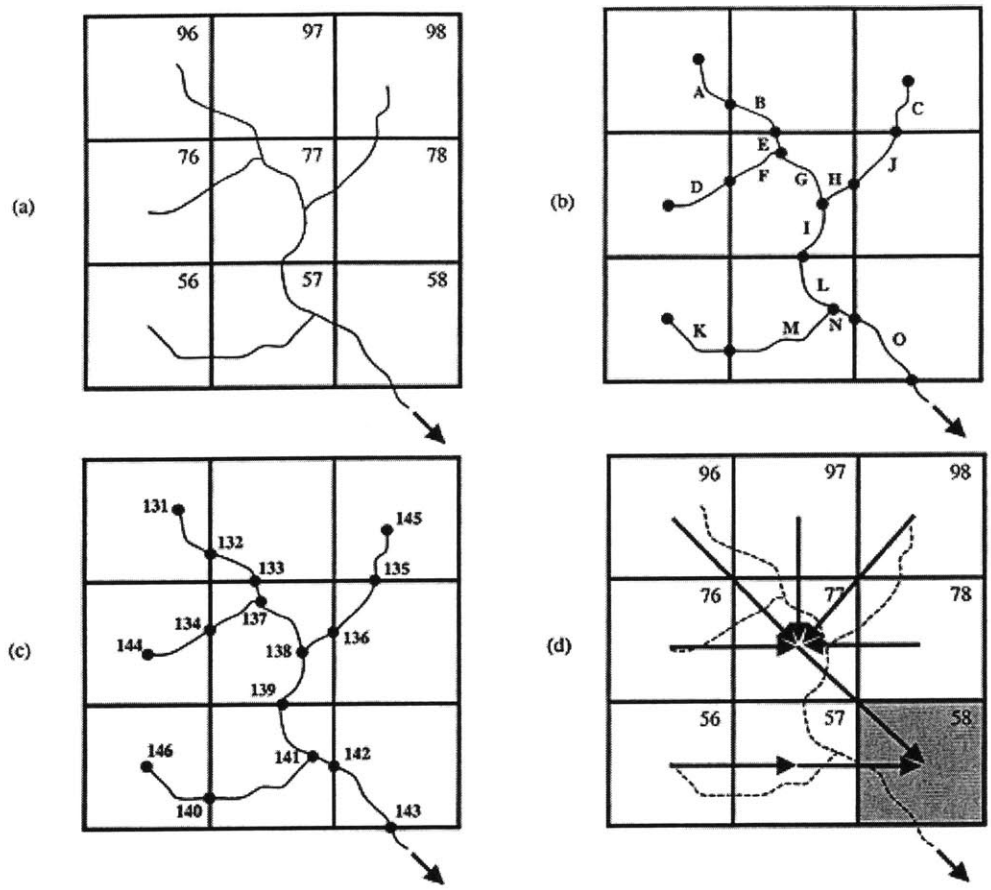


Figure 3-5: The Network Tracing Method (NTM) is a multi-step process that first, a) overlays the stream network with a coarse resolution grid, b) intersects the stream network with the grid, creating stream reaches along the network, c) creates nodes on the network for each end point, converging point, and intersected point (exit nodes) and labels the nodes numerically, and finally d) calculates the reach length from one exit node to the next to determine the upsclaed flow direction.

Operation of VB executable

The Flowrouting Visual Basic executable is composed of three modules, which must be run sequentially. The first module is the stream rank module, which takes the stream network text file as input and determines the hydrologic rank of each stream reach in the network. The stream rank module outputs a text file, which serves as the input to the next module, the upstream flow length module. As the name suggests, the upstream flow length module determines the length of the stream reach upstream to each exit node on the network. This information is used in the final part of the program, the upscaled flow direction module, to determine the direction of flow for

each grid pixel according to the criteria described in the previous paragraph. This module requires the number of columns in the grid (122 for this research), the upscaled flow direction threshold value (0.6), and the size of a grid cell side (0.5). Another input required by all three modules of the VB executable is the array size parameter. This value should be set to a value equal to or greater than the number of stream segments in the intersected network. This can be determined by looking at the length of the stream network text file. A detailed description of the modules and parameters used to run the program can be found at:
<http://ceprofs.tamu.edu/folivera/GISTools/NTMTool/ToolHelp.htm>

Flowrouting Output

The NTM VB executable outputs a significant amount of information, including the downstream pixel for each pixel contained in the coarse resolution grid (Olivera and Raina, 2003). The header of the output gives the number of columns in the grid, the upscaled flow direction threshold, cell side, Trace Arc value (the number of stream segments traced downstream to determine reach length), and side-to-corner ratio. The side-to-corner ratio obtained for this work was 63:37.

The executable output also contains column data, but only the first two columns are of interest to this project. Column 1 is the identification number of a pixel (BoxID) and column 2 is the identification number of the pixel that is directly downstream of the pixel in column 1 (DSBoxID). The identification numbers in column 1 are not repeated as there can be only one downstream pixel to each pixel. However, a pixel can be the downstream pixel to several pixels, so the identification numbers in column 2 can repeat. Pixels that are outlets are given a downstream pixel identification number of zero. An example of the output from the VB executable is given below in Figure 3-6. For a definition of the other column values, the reader is referred to Olivera and Raina's publication (2003).

```
Number of columns = 64
Threshold = 0.600
Cell side = 0.500
TraceArc =  30
Side-to-corner ratio = 64.07 / 35.93

BoxID, DSBoxID, FDBox, Reach, Meandering, ExitLineIndex, BoxU/SLength, Exit-
BoxID, ExitBoxIndex, ShortReach, NextExitBoxID
4764, 0, 0,0.000,0.000,48732,0.774,0,0,0.000,0
4752, 4687, 8,0.577,0.816,66360,1.250,4751,3,0.470,4687
4751, 4686, 8,0.639,0.904,69078,1.720,4687,64,0.107,4686
4769, 4704, 8,1.030,1.457,57436,0.795,4705,84,0.288,4704
4795, 0, 0,0.000,0.000,81529,5.842,0,0,0.000,0
4782, 4781, 16,0.544,1.088,60365,1.061,4781,48,0.544,4780
4773, 4774, 1,0.319,0.638,48745,0.674,4774,52,0.319,0
```

Figure 3-6: Sample output from Upscaled Flowrouting VB executable.

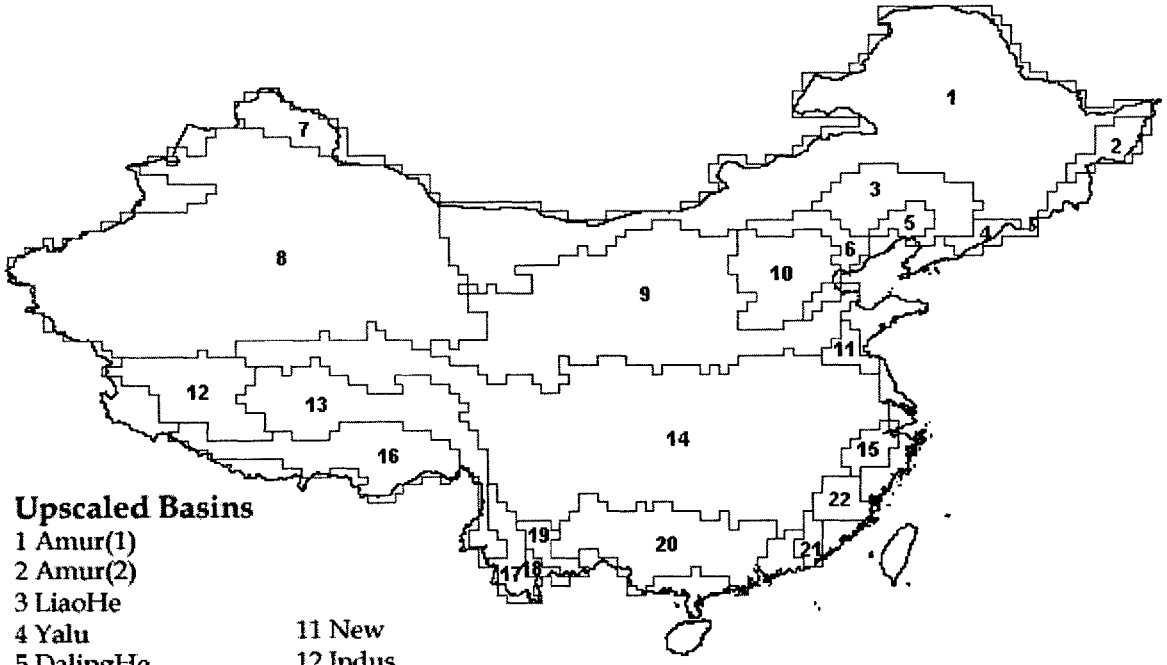
The header information, column headers, and all columns other than columns 1 and 2 are deleted and the file is saved as a tab-delimited text file. This file is read by the Pre-Processor to determine the contributing area of each pixel.

Determining Basin Boundaries

The delineation of rivers at the 0.5° x 0.5° scale was performed by taking the output from the NTM program and determining all the upstream tributary pixels, and thus the contributing area, for each pixel in the grid. To do this, first the pixels directly upstream of a pixel are found, then the upstream pixels of those pixels, and so on until all the tributary pixels are determined. The China mask grid described in Section 3.3.1 is used to filter out pixels that are not part of China while maintaining a complete network within China for basins on the nation's border.

Outlet pixels for major rivers were identified by superimposing the fine-resolution stream network on the 0.5° x 0.5° grid in ArcMap. In total, 22 major rivers were identified. The upscaled rivers have similar flow networks to those of the real rivers, however, due to the upscaling some of the resulting rivers are the combination of smaller, individual rivers. This cannot be prevented, but it does not pose a major problem for the purposes of this research. The tributary information obtained from this program for outlet pixels defines the boundaries of the upscaled basins. Other pixles not contained within one of these major river basins are considered part of a collective

river basin for water budgeting purposes, but the upscaled flow information for these pixels are maintained. Figure 3-7 shows the delineation of the upscaled river basins.



Upscaled Basins

1 Amur(1)
2 Amur(2)
3 LiaoHe
4 Yalu
5 DalingHe
6 LuanHe
7 UlungurHe
8 Tarim
9 HuangHe (Yellow)
10 HaiHe
11 New
12 Indus
13 Nu Jiang (Salween)
14 Chang Jiang (Yangtze)
15 Fuchun Jiang
16 Yarlu Zangbu Jiang (Brahmaputra)
17 Lancang Jiang (Mekong)
18 Song Da
19 Song Koi
20 Xi/Bei Jiang (Pearl)
21 Han Jiang
22 Min Jiang

Figure 3-7: Upscaled river basins of the optimization model. All pixels not defined by one of the 22 river basins are part of a collective basin in which upscaled flowrouting information is used to determine water movement.

Tributary & Basin Files

Information from the NTM is used to write four files, according to GAMS syntax, so that the tributary pixels of each pixel and each basin are known. These are the trib_def, trib_table, basin_def, and basin_table text files. The data from these files are used as inputs to the GAMS model and are used directly in the pixel water balance of Equation 3.13 (subscript n for tributary pixels) and embedded in the basin water balance. These files also indirectly define other related inputs, such as the number of pixels in a basin, B_b, and the number of tributary pixels to each pixel, T_i.

3.3.3 Pre-Processor

The Pre-Processor, written in MATLAB, was developed to perform many of the calculations that would otherwise need to be done in GAMS. GAMS has computational limitations that necessitate minimizing the number of calculations performed by the optimization model. In addition, GAMS requires that input data be of a certain format, so the Pre-Processor is also used to write the formatted input files for GAMS.

Calculating Annual Actual Evapotranspiration

Overview of Evapotranspiration

Evapotranspiration is the combination of two processes: evaporation and transpiration. These two processes convert liquid water to water vapor, a critical component of the hydrologic cycle. However, evaporation occurs from water surfaces while transpiration occurs inside plants. Evaporation is driven by the condition of the atmosphere. High temperature, low atmospheric pressure and low humidity are conditions that create high potential for evaporation. Transpiration is driven by the metabolic needs of the plant, which are also affected by the climatic conditions. Unfortunately for the scientific community, these processes occur simultaneously and are difficult to quantitatively distinguish from each other (Allen et al., 1998). Since it is difficult to separate these two processes this paper will collectively refer to transpiration and evaporation as evapotranspiration, keeping in mind that this process includes vegetation transpiration as well as evaporation from soils and surfaces.

Potential evapotranspiration (PET) is a measure of the atmosphere's ability to evapotranspire water and is determined by the climate of the region. PET data is reported based on the evapotranspiration rate of a reference crop of an idealized grass of specified height and roughness and with access to a plentiful supply of water (Doorenbos and Pruitt, 1977). Actual evapotranspiration (AET) is the amount of water

that actually evapotranspires from the surface into the atmosphere and is determined by both the climate and the land cover (e.g forest, crop, urban area).

Actual Evapotranspiration

All land cover types evapotranspire. The amount of evapotranspiration depends on the amount of water available for evapotranspiration, the potential for evapotranspiration, and the land cover type. The potential evapotranspiration for some land covers may be greater than the available water, and in these cases actual evapotranspiration cannot exceed the available water. See Figure 3-8.

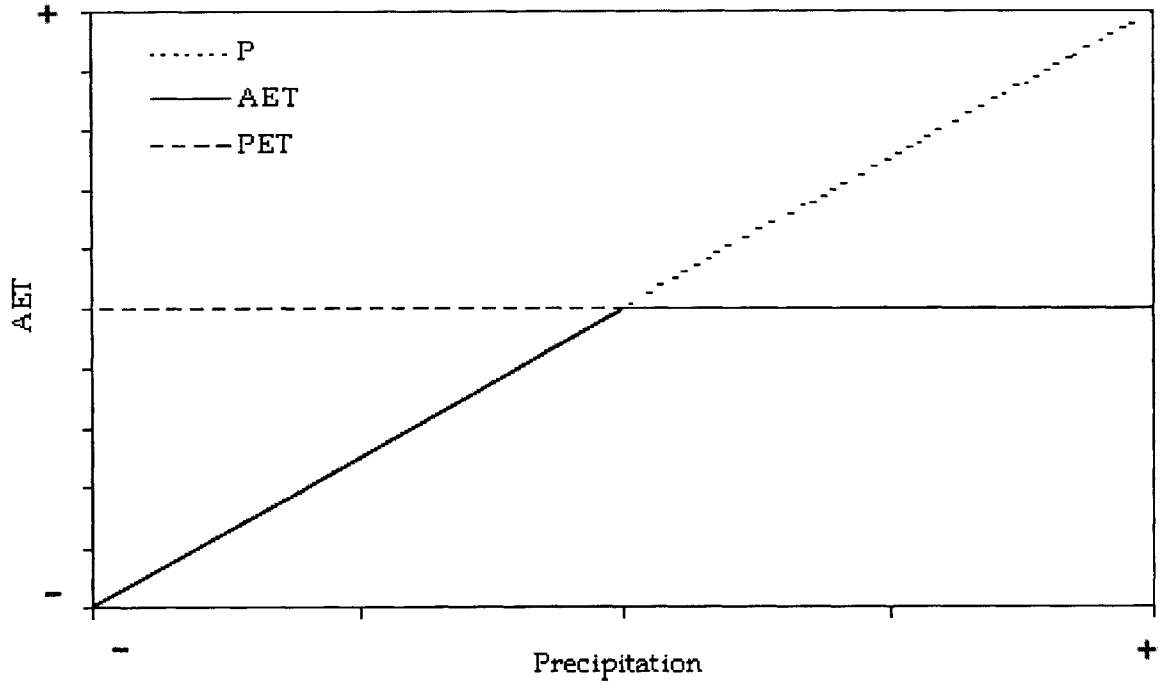


Figure 3-8: Annual actual evapotranspiration (AET) is limited by the availability of water (annual precipitation, P) and annual potential evapotranspiration (PET).

For the purposes of this model, precipitation and irrigated water are the only sources of available water since changes in storage are considered negligible. As discussed in Chapter 2, crops require water for growth and producing fruit. If a plant does not acquire the required amount of water it may become stressed, reducing its yield. Therefore, the model determines whether or not adequate water can be provided, either with precipitation or irrigation, throughout a crop's growth period. If

the water is not available the model will not choose to plant the crop. The model assumes that if a crop is grown it produces optimal yield.

Plants evapotranspire at different rates depending on the type of plant and the climate where it is grown. One acre of trees will evapotranspire more water than one acre of grass planted in the same space. Similarly, one acre of tress planted in New England will evapotranspire less water than one acre of trees planted in Florida. Evapotranspiration depends on the spatial distribution of the climate as well as the vegetation. As discussed previously, PET is reported for an idealized grass. Not all land cover types evapotranspire like an idealized grass, so crop factors are used to adjust the potential to actual evapotranspiration, or crop water requirement, with the crop coefficient approach (Allen et al., 1998),

$$AET_c = (PET)(K_c) \tag{3.21}$$

where K_c is the crop coefficient for crop c. PET is a function of climate and as such it changes with time. Similarly, plants have different developmental stages that require different amounts of water. The middle stage of development, when the majority of growth occurs, is generally the most water consumptive stage while the initial and latter stages of development typically require less water. These differences in water demands during crop development make K_c time-dependent as well. Therefore, AET_c must be determined based on time-varying PET and K_c data.

AET must also be calculated for areas other than cropland, such as forest, pasturelands, and urban areas. Factors for adjusting PET to AET can also be used for these other land use types. Since non-arable is never irrigated, this land can only evapotranspire water that is available through precipitation. As a result, AET from non-arable land is not allowed to exceed the precipitation rate (refer to Figure 3-8).

The primary responsibility of the Pre-Processor is to calculate annual actual evapotranspiration rates to be used as inputs ($E\_I_{i,s}$, $E\_D_{i,s}$) in the GAMS annual water balance constraints. The first step in calculating annual AET from cropland is to determine when and where crops can be grown. The crop sequence concept discussed in Chapter 2 is used in the Pre-Processor to define possible combinations of crops to be

grown on the land. Next, the Pre-Processor must determine when each crop of a particular crop sequence can be grown in any given pixel. This is the crop growing period. The model employs the Growing Degree Days Method for determining the growing period of crops in every crop sequence.

Growing Degree Days Method

An agronomic concept known as Growing Degree Days (GDD) is used to determine when a crop can be planted and when it should be harvested. The idea is simple; a crop must be planted after a critical temperature threshold is met so that the crop is not stressed by cold air and the plant must grow until it has accumulated enough "heat units" to reach maturity. This method is more robust than the traditional "days to maturity" method, which assumes a uniform growing period length with no regard for climatic differences. A corn plant will reach maturity faster if it is grown in southern China as opposed to northern China, so the degree days method is used in the model because it allows for spatially-varied growing period lengths while the days to maturity method does not.

Degree Days, or heat units, are accumulated by taking the average daily temperature and subtracting the temperature threshold value specific to each crop.

$$DD_t = T_{ave} - T_{threshold,c} \tag{3.22}$$

where DD_t is the degree days accumulated in day t, $T_{threshold,c}$ is the critical temperature threshold of crop c, and T_{ave} is the average daily temperature for that day. However, if the average daily temperature exceeds a specified upper threshold value, the plant does not acquire any additional Degree Days above the upper threshold. If the temperature drops below $T_{thresold,c}$ at any point during the growth period the crop is considered stressed and the Pre-Processor does not allow the sequence to be grown in that pixel. The Degree Days are summed over time and once enough Degree Days have accumulated, the crop has reached maturity and is harvested. This constraint is expressed in Equation 3.23:

$$\sum_{t=1}^{T} DD_t \geq DD_{req,c} \tag{3.23}$$

where T represents the total number of days in the growth period and $DD_{req,c}$ is the required number of degree days for crop c to reach maturity. Figure 3-9 provides a visual representation of the Growing Degree Days concept.

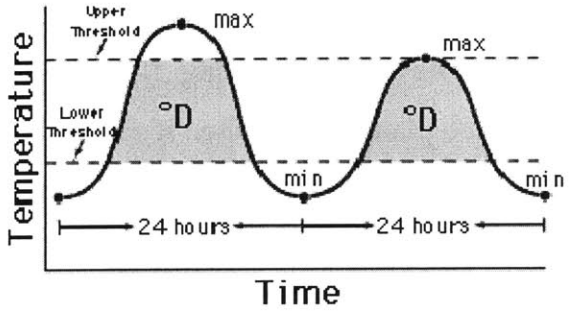


Figure 3-9: Growing Degree Days (GDD) accumulation over time. Areas in gray between the upper and lower threshold temperature values are the accrued degree days. Source: http://axp.ipm.ucdavis.edu/WEATHER/ddconcepts.html

The crops simulated in the model are defined in Section 4.4, but one crop, winter wheat, is grown differently than the other crops considered by the model. Typically, a farmer plants crops as soon as it is warm enough to grow the crop, barring any unforeseen frost events. However, winter wheat is planted in the fall and harvest in late spring. Using just the GDD methodology for determining the growth period for each crop, the Pre-Processor would choose to plant winter wheat in the spring rather than the fall. So an additional constraint is placed on the growth period for winter wheat within the Pre-Processor specifying that it cannot be planted until September. This date is consistent with winter wheat phenology information provided by Cui et al. (1984).

The number of Degree Days required by each crop to reach maturity, $DD_{req,c}$, is based on experimental observations of crop growth. The $DD_{req,c}$ and $T_{threshold,c}$ data are discussed in Section 4.4.4 of the following Chapter.

Viability of Crop Sequences

Not all 301 crop sequences can be grown in every grid pixel. In fact, many pixels in China can only support one crop per year due to the temperature limitations of the crops. An additional parameter specifying which crop sequences are viable in each pixel of the grid is written by the Pre-Processor; a sequence is given the value 1 if it is viable and 0 if it is not viable. So pixels where the temperature drop below $T_{threshold,c}$ during the crop growth period are given a viable value of 0.

Crop AET Calculation

The Degree Days requirement is used by the Pre-Processor to determine when and where crops can be grown. AET is calculated based on these crop growth periods. AET for sequences that are not viable cannot be calculated and are given the default value, zero. All viable crop sequence AET rates from irrigated land are calculated with to the crop coefficient approach according to Equation 3.24,

$$E\_I_{i,s,c} = (PE_i)(K_c) \tag{3.24}$$

where $E\_I_{i,s,c}$ is the irrigated evapotranspiration rate of crop c of sequence s in pixel i, PE_i is the PET in pixel i, and K_c is the crop evapotranspiration coefficient. The distinction between irrigated AET values and dryland AET values will be made clear later in this section. Aside from being crop specific, K_c varies throughout the stages of crop growth. Therefore the fraction of time spent in each growth stage is used to determine the appropriate K_c value for each day of the growing period calculated from the Degree Days Method. The data for K_c and fractional time in each growth stage will be presented in the following Chapter, Section 4.1.3.

Since the crop growth period and K_c both have daily temporal resolution, monthly PET data is linearly interpolated to daily PET values in order to calculate daily AET values for the crop growth period. To obtain the total AET for all crops grown in a crop sequence, Equation 3.24 is summed over time for each crop:

$$E\_I_{i,s} = \sum_{c=1}^{N_s}\left(\sum_{t=1}^{T_c}(PE_{i,t})(K_{c,t})\right) \tag{3.25}$$

where N_s is the number of crops in sequence s, T_c is the length of the growing period for crop c (determined with the GDD method), $PE_{i,t}$ is the potential evapotranspiration in pixel i for day t, and $K_{c,t}$ is the crop coefficient for crop c during the growth stage at time t.

AET for Fallow and Non-Arable Land

AET for a crop is calculated based on the climatic potential for transpiration, PET, and the water demand for the stage of development the crop is in, K_c. When a crop is not being grown on arable land, the model assumes the land is fallowed and the evapotranspiration rate is equivalent to that of bare soil. However, a water requirement does not need to be met for fallowed or non-arable land. The calculation for non-arable land and fallow land are the same, following the relation:

If precipitation at time t exceeds the PET, AET follows a similar form as Equation 3.24

$$E\_L_{i,t} = (PE_{i,t})(K_L) \tag{3.26}$$

where $E\_L_{i,t}$ is the actual evapotranspiration rate of land use L (either fallow of non-arable) in pixel i on day t, K_L is the coefficient used to adjust the potential evapotranspiration to the AET of land use type L.

If PET exceeds precipitation,

$$E\_L_{i,t} = (P_{i,t}) \tag{3.27}$$

where $P_{i,t}$ is the precipitation rate in pixel i on day t. In this way, the rate of AET never exceeds the available water, precipitation. If precipitation is high enough, land will evapotranspire at its potential but it can never evapotranspire more than the potential (refer to figure 3-8). Annual AET for non-arable land is achieved by evaluating the criteria for equations 3.26 and 3.27 and summing the resulting $E\_L_{i,t}$ values for all days in the year.

$$E\_I_{i,na} = E\_D_{i,na} = \sum_{t=1}^{365} E\_L_{i,t} \tag{3.28}$$

To arrive at the total annual actual evapotranspiration rate of crop sequence s in pixel i, including fallow periods, we use Equation 3.29,

$$E\_I_{i,s} = \sum_{c=1}^{N_s}\left(\sum_{t=1}^{T_c}\left(PE_{i,t}\right)\left(K_{c,t}\right)\right) + \sum_{t=1}^{T_F} E\_L_{i,t} \tag{3.29}$$

where T_F defines the fallow period for crop sequence s. The results of Equation 3.29 are used as inputs to the GAMS optimization model.

The degree day requirement is often met by more than one time period. Therefore the model assumes that crops are grown as soon as it becomes warm enough to grow them. Analysis of this criterio shows that modeled crop growth periods are similar to the agronomic practice of farmers (J. M. Herrmann, personal communication, May 3, 2004).

Irrigated vs. Dryland Agriculture

Arable land is divided into two categories, irrigated cropland and dryland cropland. For both land uses, crops can only be grown if the crop water requirement is met during all stages of development. Dryland cropland is distinguished from irrigated cropland in that the water requirement must be met entirely by precipitation. For dryland agriculture to be possible, precipitation must exceed the crop AET or the crop will become stressed and produce poor yield. In order for dryland crop sequences to be viable, the crops may need to be grown at different times of year than they would otherwise be grown on irrigated land. When determining the growth period for dryland crops, the Pre-Processor searches for periods when the crop water requirement will be met in addition to the temperature constraints mentioned previously in the Section on the Degree Days Method. Since the growth periods for dryland agriculture may be different from irrigated agriculture the AET values may also be different. Therefore AET for dryland agriculture must be calculated separately from irrigated agriculture.

The daily interpolation of precipitation and PET cannot be used to determine if a grid pixel is capable of supporting dryland agriculture since crops do not necessarily

need to be watered daily. Soil around plant roots retain water for a period of time that can sustain growth despite application of water. For simplicity, the model assumes that the water demand for each growth stage (initial, developing, mid, and late) must be met by precipitation. This assumption alleviates the need to know how long each crop can go between rain events to avoid stress, which is a complicated issue and beyond the scope of this research. The actual evapotranspiration rate from dryland areas is the same as that of irrigated land, but the reader should keep in mind that the indices over time can be different from irrigated to dryland crops.

$$E\_D_{i,s} = \sum_{c=1}^{N_s}\left(\sum_{t=1}^{T_c}\left(PE_{i,t}\right)\left(K_{c,t}\right)\right) + \sum_{t=1}^{T_F} E\_L_{i,t} \tag{3.30}$$

Dryland agriculture is further constrained so that the water demand of each stage of crop development is met. The water demand criteria for dryland agriculture in a given pixel is expressed in Equations 3.31 and 3.32,

$$E\_D_{i,s,c,S} = \sum_{t=1}^{D_S}\left(PE_{i,t}\right)\left(K_{c,t}\right) \tag{3.31}$$

$$E\_D_{i,s,c,S} \le \sum_{t=1}^{D_S} P_{i,t} \tag{3.32}$$

where $E\_D_{i,s,c,S}$ is the actual dryland evapotranspiration during developmental stage S of crop c of sequence s in pixel i, D_S is the number of days crop c spends in developmental stage S, and $P_{i,t}$ represents the precipitation in pixel i on day t. If this condition is not met for one or more of the developmental stages of a crop in a given crop sequence, that sequence is not a viable dryland sequence in that pixel. A text file specifying which crop sequences are viable for dryland agriculture in each pixel is written by the Pre-Processor. Therefore a viable text file is written for irrigated crop sequences (constrained only by temperature) as well as dryland crop sequences (constrained by temperature and precipitation). Logically, in areas that are water scarce more sequences are viable if irrigated water is supplied.

Writing Input Files for Optimization Model

The tributary and viable text files are not the only files written for use in the GAMS optimization model. Annual AET calculations for arable land and non-arable land, annual precipitation, pixel area, arable land area (cropland), irrigated land area, and crop yields for each crop sequence are all inputs of the optimization model and as such must be formatted to fit GAMS syntax. For all files, only pixels contained within the mask grid file are output to GAMS. This further reduces the computational burden of the optimization model.

Chapter 4

Data Discussion

The validity of any model depends not only on the conceptual framework behind the model but also on the quality of the data input to it. As mentioned in the previous chapter, the physical factors influencing agricultural production are arable land, water, climate, and type of crops grown. These are the optimization model parameters. The details of the data sources used to describe these parameters are discussed in this chapter.

4.1 Water

Plants require water during all stages of development to reach maturity so water must be made available to crops through the duration of development, either in the form of irrigation or precipitation. The data sources and concepts used to determine the amount of water available to crops are discussed in this section.

4.1.1 Precipitation

Irrigation water comes from either groundwater or surface water systems. As described in the previous chapter, precipitation feeds the groundwater and surface water systems. Therefore the original source of irrigation water is precipitation, whether it falls locally or upstream of the location the water is withdrawn. There are many data sets that describe the precipitation over China. Three different sources of precipitation data are analyzed here: the Climate Research Unit (CRU) Global Climate Dataset (New et al., 1999), TRMM product

3B43 (DAAC, 2003), and Willmott and Matsuura's monthly and annual climatology (2001a). The following sections will describe these datasets.

Climatic Research Unit (CRU) Global Climate Dataset

The Climatic Research Unit Global Climate Dataset (New et al., 1999, henceforth CRU) is a mean monthly surface climate dataset with global extent, representing the period 1961-1990 with a $0.5^{o} \times 0.5^{o}$ spatial resolution. It contains data for precipitation, wet days, temperature, diurnal temperature range, vapor pressure, relative humidity, sunshine, cloud cover, ground frost days, air frost days, and wind speed. The dataset is the result of combining data from many meteorological stations; 19,295 rain gage stations contributed to the precipitation climatology.

The station data was interpolated to a global grid using thin-plate splines. This method treats climatic variables as functions of latitude, longitude and elevation. The National Meteorological Agencies supplied 44 percent of the global precipitation data. Additional sources of precipitation data were the World Meteorological Organisation (WMO, 1996), previous CRU datasets (Eischeid, 1991), the Centro Internacional de Agricultura Tropical (CIAT), Müller (1982), FAO (1984), and the U.S. Air Force Climatological Data Volume (USAF, 1987). An extensive quality control check was performed before the data was interpolated to a global data set (New et al., 1999).

Figure 4-2a shows the CRU annual climatology for China.

TRMM

NASA and the Japan National Space Development Agency (NASDA) launched the TRMM satellite in November 1997. It covers the global latitude belt between 38^{o}N and 38^{o}S and was designed with a 3-year mission life. The TRMM satellite measures and collects visible and infrared radiance, microwave brightness temperature, radar reflectivity, surface radar volume scans, and rain gage and disdrometer data. The TRMM Science Team has developed many

algorithms to process the raw data and convert them to useful climate datasets (DAAC, 2003). TRMM data is distributed by the GES Distributed Active Archive Center (DAAC).

TRMM product 3B43 is the second source of precipitation considered for this model. It is derived from four other TRMM precipitation products: the monthly average unclipped TRMM Microwave Imager (TMI) estimate, the monthly average Special Sensor Microwave/Imager (SSM/I) estimate, the pentad-average adjusted merged-infrared (IR) estimate, and the monthly-accumulated Climate Assessment and Monitoring System (CAMS) or the Global Precipitation Climatology Centre (GPCC) rain gage analysis. The IR product was derived from remotely sensed IR estimates using the combined instrument rain calibration algorithm 3B42. The TMI product is an intermediate product of algorithm 3B42. The SSM/I product was developed with TRMM algorithm 3A46. Rain gage products from CAMS and GPCC were processed with algorithm 3A45. The final TRMM product 3B43 (henceforth TRMM) extends globally from 40°N to 40°S, with a resolution of 1° x 1°. The coverage period analyzed in this paper is 1998 through 2002 (DAAC, 2003).

Figure 4-2b shows the TRMM annual climatology for China.

Willmott and Matsuura (W&M)

The Willmott and Matsuura (2001a) Terrestrial Air Temperature and Precipitation Climatology, Version 3.02, an average monthly gridded climatological dataset of precipitation and air temperature. The product is a global climatology covering the period 1950 through 1999, with a spatial resolution of 0.5° x 0.5°. This dataset (henceforth W&M) was produced by combining station data from the Global Historical Climatology Network (GHCN version 2) with station data from Legates and Willmott (1990a and b). Globally, 20,599 GHCN stations were used for precipitation; 26,858 stations were used for

precipitation from the Legates and Willmott data archive (Willmott and Matsuura, 2001a).

The station data were converted to a gridded format using Climatologically Aided Interpolation (CAI) and traditional interpolation. CAI was performed to derive the difference between station monthly values and station monthly climatological average values. Traditional interpolation is then used on the station differences to derive a gridded difference field. Cross validation was performed to estimate interpolation errors, which are also provided in gridded format (Willmott and Matsuura, 2001a). Figure 4-2c shows the W&M annual climatology for China. The reader should note the disparity in the scales of Figure 4-2.

Discussion

The three sources of precipitation data all have different temporal and spatial resolution. All 3 sources provide average monthly precipitation. However, the TRMM dataset has the shortest coverage period of just 5 years. The W&M climatology has the longest coverage period, 50 years, which makes it least susceptible to the affect of individual anomalies. Since it is a long-term average, the W&M data is the most appropriate for representing a typical year, as described in Chapter 2. The CRU and W&M datasets have 0.5^{o} x 0.5^{o} spatial resolution, which is twice the resolution of the TRMM dataset. Figure 4-1 shows the average monthly precipitation rate for all three sources over China up to 40^{o}N (extent of TRMM data). Table 4.1 shows the total annual volume of precipitation for all three sources over all of China up to 40^{o}N.

Precipitation Source	Annual Volume of Water [km3]
TRMM	7,071
W&M	4,949
CRU	1,599

Table 4.1: Comparison of TRMM, Willmott and Matsuura and Climate Research Unit annual volumetric precipitation over China.

TRMM data gives the highest estimate of precipitation while CRU gives the lowest estimate. TRMM is more than 320% higher than CRU data and 40% higher than the W&M precipitation data. W&M is 200% greater than CRU. The range of estimates is considerably large.

With the best coverage period, highest available resolution, and middle of the range estimate, the W&M precipitation climatology was selected for use in the model.

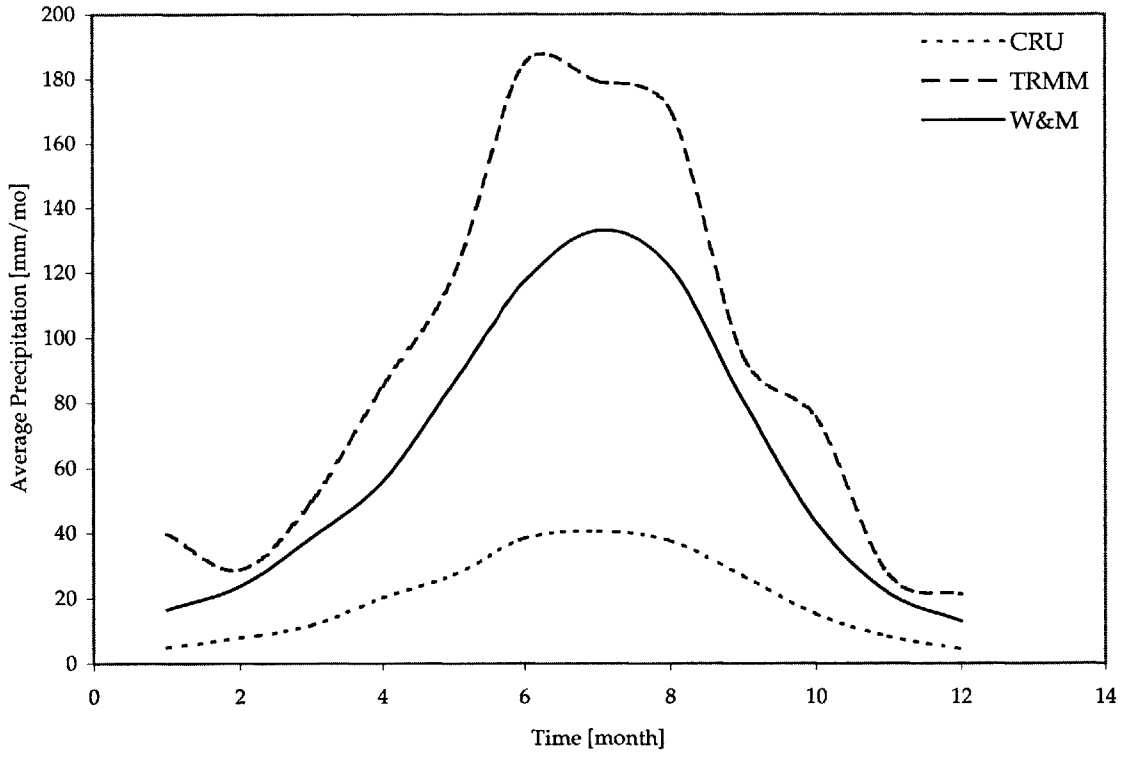


Figure 4-1: Annual variation of average monthly precipitation across China up to 40°N.

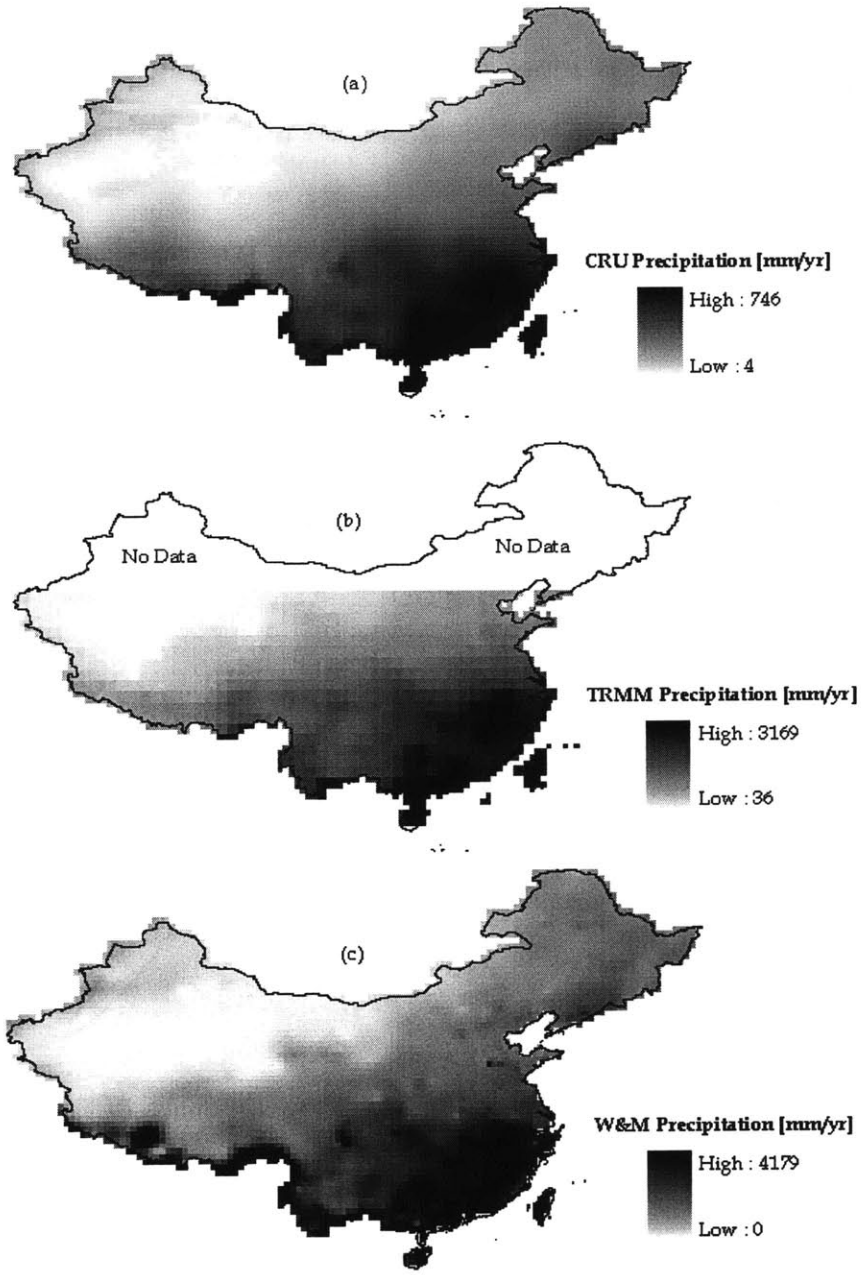


Figure 4-2: Annual precipitation rates from (a) CRU, (b) TRMM, and (c) Willmott & Matsuura datasets. Note the disparity in the scales of each data set.

4.1.2 Evapotranspiration

Evapotranspiration is one of the key components of the water balance described in Chapter 3. The distinction between AET and PET was also

explained in the previous Chapter. In this section we describe data for PET, used in the optimization model, and AET, used for comparison with the model results.

Potential Evapotranspiration

Though it is actually a measure of local climatic conditions, potential evapotranspiration is defined as the amount of water that would be evapotranspired by a reference crop, an idealized grass of specified height and roughness and with access to a plentiful supply of water (Doorenbos and Pruitt, 1977). The Willmott and Matsuura (2001b) Terrestrial Water Balance Data Archive, Version 1.02, contains a gridded climatological dataset of adjusted average monthly potential evapotranspiration. The PET data is in 0.5^{o} x 0.5^{o} spatial resolution and is a climatology covering 1950 through 1999 (Willmott and Matsuura, 2001b). Unlike precipitation, PET cannot be measured directly from a sensor. Instead, it is calculated using a modified version of the Thornthwaite water balance procedure, which derives PET from air temperature data while conserving mass (Willmott et al., 1985). While other PET datasets exist, the W&M PET data is selected for this model because it has the desired spatial resolution of 0.5^{o} x 0.5^{o} and it is consistent with the W&M precipitation dataset. Figure 4-3 shows the annual climatology of the W&M PET dataset for China.

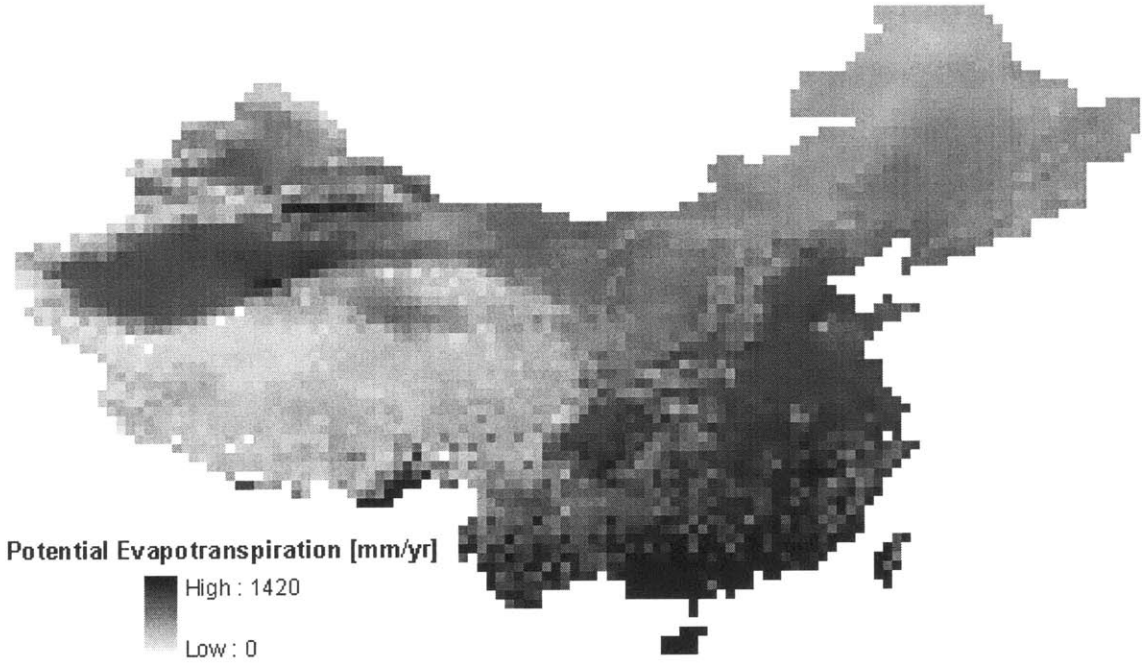


Figure 4-3: Annual potential evapotranspiration rate.

Actual Evapotranspiration

The Willmott and Matsuura (2001b) Terrestrial Water Balance Data Archive, Version 1.02, also contains a gridded climatological dataset of average monthly actual evapotranspiration. AET was calculated from a water balance based on semi-empirical relationships between measured precipitation and estimated PET, described in the previous sections. The AET data is in 0.5° x 0.5° spatial resolution and is a climatology covering 1950 through 1999 (Willmott and Matsuura, 2001b). Figure 4-4 shows the annual climatology of actual evapotranspiration in China.

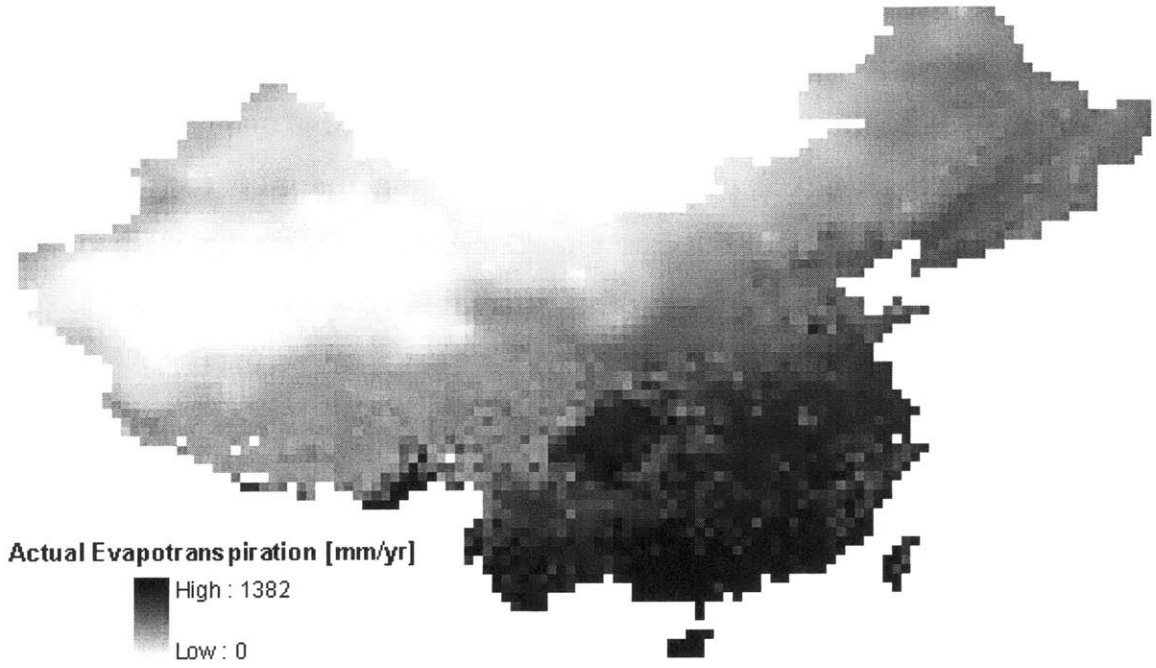


Figure 4-4: Annual actual evapotranspiration rate.

4.1.3 Evapotranspiration Adjustment Factors

In Equation 3.26 we introduced the variable K_L as the coefficient used to adjust the potential evapotranspiration to the potential of land use type L. Unfortunately, data for adjustment coefficients of land uses other than crops is severely limited. The K_L for non-arable land is assigned a value of 0.69. This value was determined by averaging the ratio of the W&M AET data to PET data. For fallowed land we assume a K_L of 0.75 since fallowed land should evapotranspire less water than cropland. The adjustment factor for fallow land should therefore, be lower than the average crop coefficient.

The evapotranspiration adjustment factors for crops are called crop coefficients and were previously mentioned throughout Section 3.3.3 during the discussion of AET calculation. The crop coefficients used in the model are summarized in Table 4-2, below. There are three sources of Kc data: Kang et al. (2003), Li et al. (2003), and Allen et al. (1998).

Crop coefficients for winter wheat are based on data from Kang et al. (2003) while coefficients for spring wheat and maize are based on Li et al. (2003). For each of these crops the coefficients were derived in the same manner. Kang et al. (2003) and Li et al. (2003) provide figures showing the crop coefficient as a function of time throughout the crop growth cycle. These figures were visually interpreted to arrive at K_c values for the different developmental stages. K_c data for rice, vegetables, tubers, and oil are derived from Allen et al. (1998). Allen et al. (1998) include K_c values for many different crops, including several crops for each general crop category (e.g. wheat). The K_c for each growth period is calculated by averaging the K_c values for each crop listed under the main crop category. The crops listed were selected based on their role in the Chinese diet and temperature sensitivity, as will be discussed in Sections 4.4.1 and 4.4.4.

	Kc		
	Initial	**Mid**	**End**
Winter Wheat[1]	0.6	1.25	0.9
Spring Wheat[2]	0.55	1.2	0.7
Rice[3]	1.05	1.2	0.75
Maize[2]	0.5	1.2	0.65
Tuber[3]	0.5	1.1	0.95
Oil[3]	0.35	1.15	0.35
Vegetable[3]	0.54	1.07	0.78

Table 4.2: Crop evapotranspiration adjustment coefficients for the seven model crop categories. Sources: 1) Kang et al. (2003), 2) Li et al. (2003), and 3) Allen et al. (1998).

The growth period lengths were used to determine the fractional time that each model category spends in each growth stage. The same sources were used to derive the growth stage fractional time values. For winter wheat, spring wheat, and maize, the figures in Kang et al. (2003) and Li et al. (2003) were visually interpreted to determine the length of each growth stage for a crop. These lengths were divided by the total length of the growth period to arrive at a fractional time for each stage of development. For rice, vegetables, tubers, and oil the length of each growth stage for all crops listed under one of these crops in

Allen et al. (1998) was averaged and divided by the total average growth period for the model crop category to obtain the growth stage fractional time. The fractional time is applied implicitly in Equation 3.25 to determine which K_c value to use for each day of the growth period. These data are summarized in Table 4.3, below.

	Fraction of Time Spent in Growth Stage			
	Initial	**Devel**	**Mid**	**Late**
Winter Wheat[1]	0.438	0.375	0.062	0.125
Spring Wheat[2]	0.2	0.167	0.333	0.3
Rice[3]	0.182	0.182	0.424	0.212
Maize[2]	0.15	0.05	0. 6	0.2
Tuber[3]	0.206	0.229	0.371	0.194
Oil[3]	0.166	0.269	0.352	0.214
Vegetable[3]	0.196	0.254	0.391	0.159

Table 4.3: Fractional time spent in each growth stage for each model crop category. Sources: 1) Kang et al. (2003), 2) Li et al. (2003), and 3) Allen et al. (1998).

4.1.4 Upscaled River Basins

Hydro1k DEM

The Hydro1k dataset is used to derive the upscaled river basins according to the methodology described in Section 3.3.2. Hydro1k is a global topographic and hydrologic dataset with 1-km horizontal resolution and 1-meter vertical resolution. It is based on GTOPO30, an elevation dataset created by the USGS from 8 different sources (LP DAAC, 2004). 1-km horizontal resolution is the highest spatial resolution for global topographic data currently available (Tom Farr, personal communication, November 2002), though higher resolution is available for some parts of the world, such as the United States.

Other hydrologic data, including flow direction, flow accumulation, slope, aspect, and compound topography index (CTI) are also provided with the Hydro1k dataset. The Hydro1k elevation data was processed to remove sinks. Flow direction and flow accumulation grids are necessary for deriving the fine

resolution stream network (as described in Section 3.3.2) and were derived from the Hydro1k elevation data instead of using the flow direction/accumulation data provided by Hydro1k as a precautionary step.

Figure 4-5 shows the Hydro1k digital elevation map (DEM) for China.

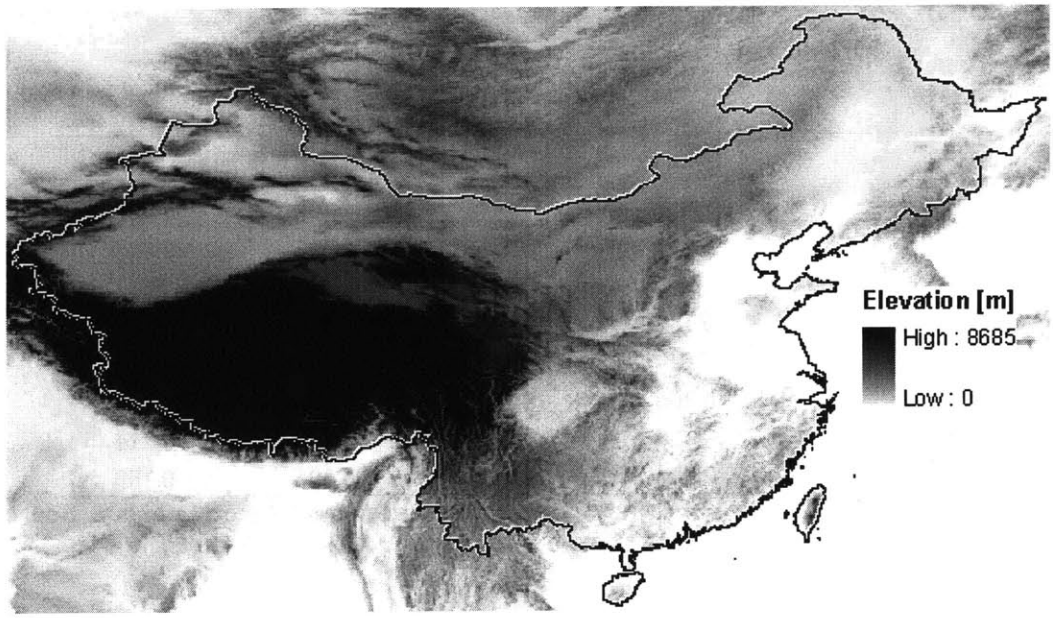


Figure 4-5: Topography of China (Source: Hydro1k DEM, LP DAAC, 2004)

Fractional Basin Population

Provincial population data from the 1990 Chinese Statistical Yearbook (State Statistics Bureau, 1990) was used to estimate F_b, the fraction of the total population residing in each upscaled basin. The Chinese county boundary coverage provided by CITAS (1997) is aggregated by provincial code to create a provincal boundary coverage. Working in Lambert Azimuthal Equal Area projection, the provincial coverage is intersected with the boundaries of the upscaled river basins to determine what fraction of each province is contained in each basin. These fractions are multiplied by the provincial populations and summarized for each basin to determine the basin populations. In this way the total Chinese population is accounted for. Fractional basin population, F_b, is

used to calculate the municipal and industrial water use in each basin according to Equation 3.11. The basin populations are presented in Appendix C.

4.2 Arable Land

There are many available data sources of land use, including remotely sensed sources and census data. The model considers two types of arable land, total cropland and irrigated cropland. Land use data can be used to determine the amount of arable land in China. Sections 4.2.1 and 4.2.2 describe the sources of land use data considered for quantifying these two land use types and gives an explanation of which data source is best for use in the optimization model.

4.2.1 Cropland

Remotely sensed land use data with global coverage is available from several space borne sensors, including the Moderate Resolution Imaging Spectroradiometer (MODIS), Advanced Very High Resolution Radiometer (AVHRR), and Landsat Thematic Mapper (Landsat TM). There are many differences between these three land use sources and those differences are explored in the following sections. In addition, a ground-based census of cropland is used for comparison and validation of the remotely sensed sources.

MODIS

MODIS land cover type product, MOD 12, is a 1-km spatial resolution land cover dataset with global coverage. MODIS data is available from the Earth Observing System (EOS) Data Gateway and is provided in sinusoidal map projection. MOD 12 is intended for use with global and regional-scale models of hydrologic processes, among other uses, making it ideal for the purposes of this project (DAAC, 2004). The MODIS instrument is flown aboard the Terra and Aqua satellites. Terra and Aqua cover the Earth's surface every 1-2 days while MODIS collects data in 36 spectral bands.

The seasonal variations in MODIS spectral bands 1 through 7 are used to determine land use through two measures of vegetation, the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI). Both indices measure the land surface "greenness" by looking at the reflectance of spectral bands in the electromagnetic spectrum that are sensitive to vegetation. The calculation for NDVI is straightforward,

$$NDVI = \frac{[X_{NIR} - X_{Red}]}{[X_{NIR} + X_{Red}]} \tag{4.4}$$

where X_{NIR} is the reflectance in the near infrared spectral band and X_{Red} is the reflectance in the red spectral band. NDVI values range from –1 to +1, with positive numbers indicating increased "greenness."

The EVI uses the blue spectral band to enhance the accuracy of the index classification. EVI is calculated with the following equation,

$$EVI = G \times \frac{[X_{NIR} - X_{Red}]}{[X_{NIR} + C_1 X_{Red} - C_2 X_{Blue} + L]} \tag{4.5}$$

where X_{Blue} is the reflectance in the blue wavelength, G is the gain factor, L is a canopy background adjustment factor, and C_1 and C_2 are coefficients that correct for aerosol affects on the reflectance in the blue spectral band. G, C_1, C_2, and L are empirically determined to be 2.5, 6.0, 7.5 and 1.0, respectively (Miura et al., 2001). Similar to NDVI, higher EVI values indicate increased vegetation.

The MOD 12 land cover grid is derived with a neural network classifier, which uses 1-km topographic data and 32-day NDVI and EVI composites to conduct the land cover classification. Compositing is done to minimize cloud contamination and is achieved by taking the maximum NDVI or EVI value from a 32-day period. Land cover is classified according to the International Geosphere Biosphere Programme (IGBP) global vegetation database, which includes two categories for cropland: cropland and cropland/natural vegetation mosaic. By definition, 'Cropland/natural vegetation mosaics are lands with a mosaic of croplands, forests, shrublands, and grasslands in which no one

component comprises more than 60% of the landscape' (WRI, 2003). Given the ambiguity of this land cover classification, total cropland can be calculated under several assumptions. The first is to assume that none of the land classified as cropland/natural vegetation mosaic in a 1-km grid cell is cropland. The second choice is to assume that approximately 50% of the land classified as cropland/natural vegetation mosaic is cropland. The third choice is to assume 100% of the land classified as cropland/natural vegetation mosaic is cropland. These three assumptions result in different estimates of total arable land in China (see Table 4.4 below).

IGBP classification assumption	Total Arable Land in China [10^6 km^2]
Only cropland category	2.14
50% of cropland/natural veg. mosaic	2.29
100% of cropland/natural veg. mosaic	2.45

Table 4.4: Comparison of MODIS land cover data using different assumptions about classification category 'cropland/natural vegetation mosaic.'

These estimates were obtained by first projecting the MOD12 data from sinusoidal projection to Lambert Azimuthal Equal Area projection optimized for Asia, then clipping the extent of the data with the boundary of China using ArcInfo. The boundary of China used to clip the MODIS data is derived from the county boundary coverage supplied by the China In Time and Space (CITAS) project (CITAS, 1997). The county boundaries are merged to create a continuous coverage for all of China. By clipping the MODIS land cover data with the boundary of China, only land within the boundary of China is considered. The resulting grid is then summed over all the cells to obtain the total cropland in China (recall from Chapter 2, for the purposes of this study arable land is defined as the total current cropland). As an equal area projection, the Lambert Azimuthal projection is used because it is appropriate for measurements of area.

AVHRR

The Global Land Cover Characterization (GLCC) is a 1-km resolution land cover database derived from Advanced Very High Resolution Radiometer (AVHRR) data. 10-day NDVI composites were calculated from AVHRR data collected from April 1992 to March 1993. These composites are used in an unsupervised classification to determine land cover following the IGBP land cover classification scheme mentioned in the previous section. The GLCC database covers the entire globe and is available from the Land Processes Distributed Active Archive Center (DAAC). The Eurasia GLCC database, which contains data for all of China, is provided in Lambert Azimuthal Equal Area projection, optimized for Europe and Asia (LP DAAC, 2003).

Under the same cell area assumptions described in the MODIS land classification discussion, the total arable land for China using the GLCC AVHRR data are shown in Table 4.5, below.

IGBP classification assumption	Total Arable Land in China [106 km2]
Only cropland category	1.79
50% of cropland/natural veg. mosaic	2.26
100% of cropland/natural veg. mosaic	2.73

Table 4.5: Comparison of GLCC AVHRR land cover data using different assumptions about classification category 'cropland/natural vegetation mosaic.'

Landsat TM

Frolking et al. (2002) combined Landsat Thematic Mapper (TM) data with county-scale agricultural census data to produce 0.5° x 0.5° resolution maps of agriculture data in China. The data are provided in geographic coordinates by the University of New Hampshire (UNH) Denitrification-Decomposition (DNDC) scientific team. The Landsat TM data used to create the 0.5° maps was primarily from the years 1995-1996 and had 100-m spatial resolution. The agricultural census data are from the Eco-Environmental Database of the

Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences (Frolking et al., 2002).

The TM sensor is a multispectral scanning radiometer carried onboard the Landsat 4 and 5 satellites, beginning in 1982. TM collects images of the land surface in seven spectral bands with 30-m spatial resolution. The images used for the Frolking et al. study were visually interpreted and digitized to create thematic maps of land cover. Land cover was classified into 25 categories, including two categories for cropland: paddy fields and nonflooded cropland. In addition, a stratified, multi-layer sampling design was developed to determine the fraction of noncultivated land within polygons classified as cropland (e.g., housing, footpaths, narrow roads) Validation was conducted on the classification scheme using field surveys (Frolking et al., 2002).

The agricultural census data provided by the Eco-Environmental Database contains county-scale information on sown area for 17 major crops as well as total cropland area and total sown area, among other things. This database was compared with provincial and national production statistics and found to be within 1% agreement (Frolking et al., 2002).

The DNDC maps provide the total area used for individual crop types as well as total area for single-, double-, and triple-crop land. The total area of cropland in each 0.5° grid cell was based on an aggregation of Landsat TM data while the area for single-, double- and triple-crop land was determined using an algorithm to combine the agricultural census data with the Landsat TM cropland data. For a detailed description of this process, the reader is referred to Frolking et al. (2002).

When added together, the single-, double- and triple-crop maps combine to create a total cropland area map. At 0.5° x 0.5° resolution, this dataset is much more coarse than the two previously described datasets. Each grid cell of the map measures the total cropland in China in that grid cell. The total cropland area in China, achieved by summing the total cropland area over all the grid cells

in the DNDC derived map, is 1.32 million km^2. This figure is 27% lower than the lowest estimate from either of the remotely sensed land cover data sources.

Table 4-6 summarizes all the estimates of cropland obtained from these three sources.

Source	Classification	Total Cropland in China [10^6 km^2]
AVHRR	Only cropland category	1.79
	50% of cropland/natural veg. mosaic	2.26
	100% of cropland/natural veg. mosaic	2.73
MODIS	Only cropland category	2.14
	50% of cropland/natural veg. mosaic	2.29
	100% of cropland/natural veg. mosaic	2.45
UNH Landsat TM		1.32

Table 4.6: Summary of cropland data source estimates.

Agricultural Census

The Chinese Statistical Bureau issues an annual yearbook detailing the production, growth, and other characteristics of China. Among the facts and figures are statistics on cultivated land. The 2002 edition of the yearbook reports the total cultivated land in China to be 1.3 million km^2. This figure is based on a 1996 survey, conducted each decade, of cropland usage in the country.

This estimate of 1.3 million km^2 is most likely an underestimate of the total cropland in China. Individual farmers appear more productive if they produce the same amount of food on less land, so it is in their best interest to under report their total cultivated land. Still, this is the only ground-truth estimate of cropland in China and as such it is the basis of comparison for all other sources of cropland data.

Discussion

The lowest MODIS estimate of cropland in China is 2.14 million km^2 and the lowest AVHRR estimate of cropland in China is 1.79 million km^2, while the

Landsat TM based estimate is 1.32 million km^2. Landsat TM data was taken around the same time period the cropland census was conducted, making these two sources temporally consistent. In addition, the original Landsat TM data has the highest resolution of all the remotely sensed data sources (30-m), making it more accurate at a finer scale and therefore less likely to be over or underestimating cropland. Based on these inherent strengths, the Landsat TM cropland data is used as the cropland area in the model. However, sensitivity analysis is conducted using the MODIS and AVHRR estimates as upper bound limits on arable land. Figure 4-6 shows the Landsat TM derived cropland dataset used in the model.

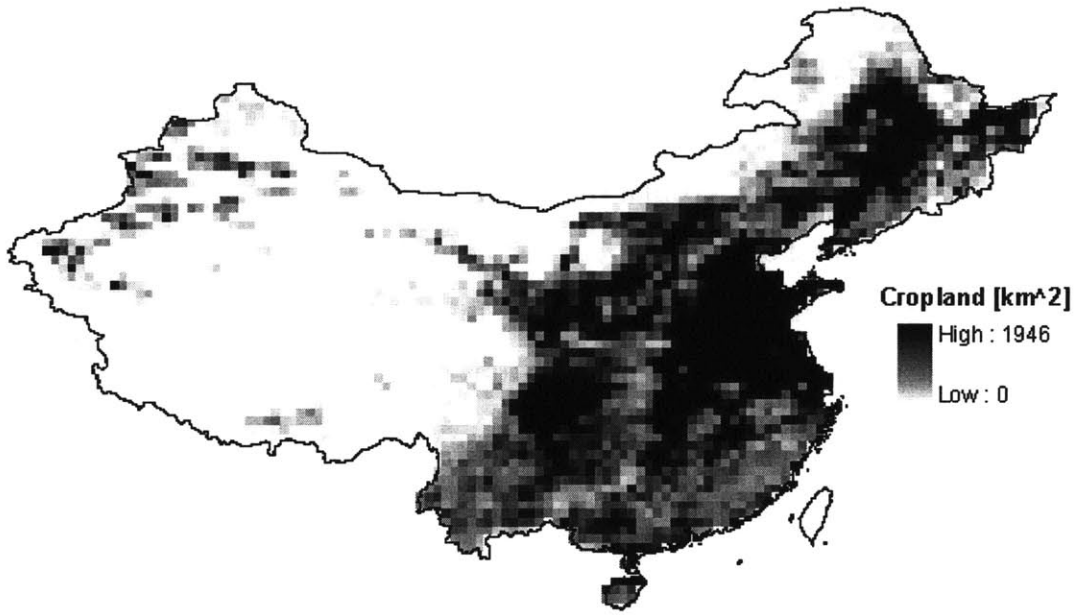


Figure 4-6: Cropland distribution within China
(Source: UNH Landsat TM derived cropland data)

4.2.2 Irrigated Cropland

Döll & Siebert

A global map of irrigated area, in 0.5° x 0.5° spatial resolution, is available from Döll and Siebert (1999). The map was generated 'for the purpose of global modeling of water use and crop production,' (Döll, 2004). It provides the total

area in each grid cell equipped for irrigation in 1995. The actual irrigated area is somewhat smaller, as not all land equipped for irrigation is used. Data for the map is based on maps outlining irrigated areas, FAO data on total irrigated area in a country, and county-, federal state- and drainage basin-scale irrigated area data, where available. For China, county-scale irrigated area data was available from the Statistics Publishing House (1991). However, this data set was missing data for 25 counties so Döll and Siebert assumed there was no irrigation in these counties. All the data sources were digitized and aggregated to the 0.5^{o} x 0.5^{o} level. Figure 4-7a shows the distribution of irrigable land in China.

From this map the total irrigated area in China is 0.46 million km^2 (Döll and Siebert, 1999). Comparing this figure to the DNDC cropland data, we find that approximately 35% of China's cropland is irrigated. However, making a pixel-to-pixel comparison of cropland to irrigated land we found that some pixels had more irrigated land than cropland. Since irrigated land should not exceed cropland, the irrigated data is filtered so that any pixels where the irrigated area exceeded cropland area are set equal to the cropland area. This creates an irrigation dataset that is consistent with the cropland dataset and is shown in Figure 4-7b.

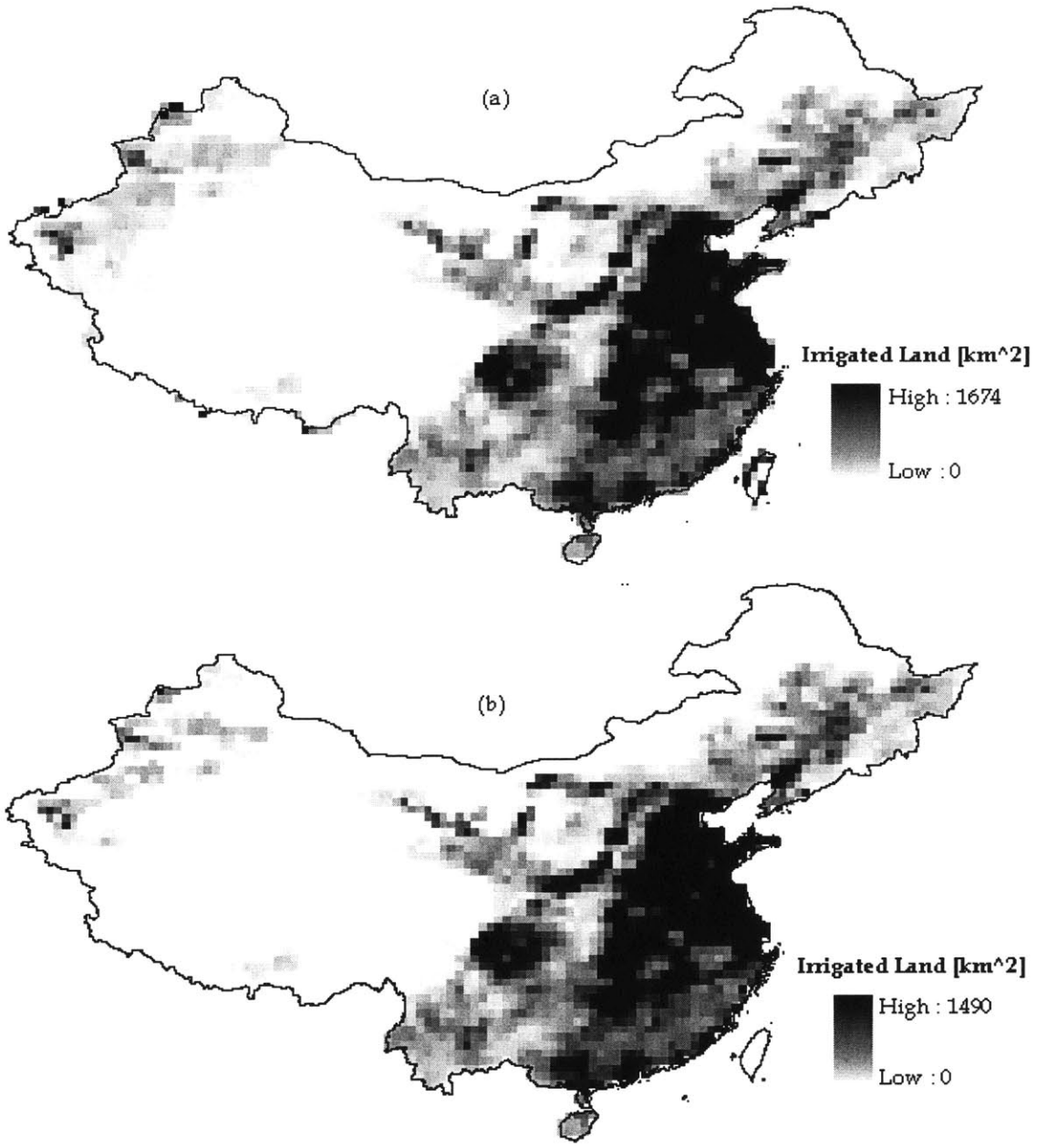


Figure 4-7: Maps of irrigated land in China, (a) Döll and Siebert original data, (b) irrigated area consistent with UNH Landsat TM derived cropland data.

4.3 Temperature

Temperature is used in the model to determine the time of year crops can be grown due to the temperature sensitivity of individual crop types. The Willmott and Matsuura Terrestrial Air Temperature and Precipitation climatology contains average monthly air temperature data. The temperature data is

provided in degrees Celsius at 0.5° x 0.5° spatial resolution and covers the same time period as the rest of the climatology, 1950-1990. As described previously, the dataset was produced by combining station data from GHCN version 2 with station data from Legates and Willmott (1990a and b). Globally, 7,280 air temperature GHCN stations were used; 24,941 stations were used for air temperature from the Legates and Willmott data archive. The station data were converted to a gridded format using Climatologically Aided Interpolation (CAI) (Willmott and Matsuura, 2001a).

4.4 Crop Type

The choice of crops grown in the model affects the food and water balance, as well as when crops can be planted. As described in section 4.1.2, crop type affects the AET portion of the water balance. Sections 4.4.1-4.4.3 will explain how the crop type affects production and consumption in the model. Section 4.4.4 will describe the agronomic thresholds for crop growth.

4.4.1 Crop Selection – Dietary Component

The model needs to produce enough food types to adequately describe the typical Chinese diet while simultaneously keeping within the limitations of computational ability of the computer the model is run on. Each additional crop adds many more new variables in the model, so crop selection is a non-trivial task.

The FAO provides a statistical database called FAOSTAT (2004a), which provides country specific nutritional information. Food consumption is included on annual food balance sheets, including the daily caloric consumption of many crop categories. The entire daily caloric consumption is calculated by summing the daily caloric consumption from each food group listed on the food balance sheets. The largest food groups, wheat, rice, maize, starchy roots, oil crops,

vegetable oils, vegetables, meats and animal fats, comprise at least 90% of the daily caloric demand for each year from 1995-2001. The food balance sheets also include daily protein and fat consumption from the crops. Summing over the same food groups, 87% of protein and 97% of fat consumption is accounted for by these categories in each year from 1995-2001. A summary of these calculations is available in Appendix B.

The model is intended to simulate agricultural production, but the Chinese diet also includes meat. This means crop production must account for the production of animal meats and products. Enough of each crop must be produced to fulfill the demand for direct human consumption as well as feed for livestock. This assumption embeds the production of animal products in crop production. As a result, the model crops are wheat, rice, maize, starchy roots (tubers), vegetables, and oil crops (including oil and vegetable oil crops). In this model, meat produced from grazing is neglected both in the diet and in production.

4.4.2 Crop Consumption – Dietary Demand

As described in the previous section, each crop considered in the model makes up a portion of the typical Chinese diet. The model must grow enough of each crop to satisfy the food demanded by the population. The demand for each crop is calculated as:

$$D_c = (N)(C_c) \tag{4.1}$$

where D_c represents total demand in units of mass per annum, N is population in units of people, and C_c is consumption in units of mass per person per annum. In section 4.4.1 the dietary demand was determined by caloric intake of crops, since diet is determined by the calories humans consume, not the mass consumed. However, production is measured in terms of mass, not calories, so for the food balance it is easier to think of consumption in terms of mass consumed.

FAOSTAT food balance sheets (2004a) also include supply and domestic utilization figures for many crops. Domestic utilization includes crop use for animal feed, seed, food manufacturing, waste, other uses, and direct human consumption. While the model is only concerned with human consumption (food manufacturing and direct human consumption), similar to animal feed, the other categories of domestic use must be taken into account in the diet in order to provide realistic estimates of the water and land required to satisfy dietary needs.

All the domestic utilization categories are included as part of the typical Chinese diet. Data from 1995-2001 was averaged to obtain the amount of each crop category required to meet per capita domestic utilization demand. These figures are summarized in Table 4.7.

Crop	C_c [kg/person/yr]
Wheat	89.02
Rice	101.83
Maize	95.14
Tubers	145.9
Oil crops	53.46
Vegetables	211.28

Table 4.7: Amount of each crop category that need to be grown in order to satisfy dietary needs, based on domestic utilization from FAO food balance sheet data (1995-2001).

4.4.3 Production

Expected Yield

According to the food balance, the model allocates land and water to maximize the population fed. In order to allocate land the model decides how much land to use for each crop sequence. Each crop that is grown has an expected yield, assuming the plants are not stressed during development. Therefore, production is the product of the area dedicated to a crop and the expected yield of the crop, as follows:

$$P_c = (L_c)(Y_c) \tag{4.2}$$

where P_c is production of crop c in units of mass, L_c is land devoted to crop c in units of area, and Y is yield in units of mass per unit area.

The FAOSTAT data (2004b) includes an agricultural production database. This includes annual yield values, in units of mass per unit area, for most crops grown in a country. These figures are reported for each calendar year, so the crop yields from the 5 most recent years, 1999-2003, were averaged to obtain an expected yield for each crop in the model (FAOSTAT, 2004b). Table 4.8 summarizes the annual crop yield values, assumed uniform over all of China, for the 6 crop categories.

Crop	Y_c [kg/ha]
Wheat	3.83
Rice	6.21
Maize	4.81
Tubers	17.63
Oil crops	0.53
Vegetables	18.41

Table 4.8: Average crop yield factors for model crop categories, based on FAOSTAT yield values from 1999-2003 (FAOSTAT, 2004b)

Crop Imports

The model assumes that current levels of imports and exports are maintained so the crop supply information provided in the FAOSTAT food balance sheets (2004a) are used to determine the annual net import of each crop category. Annual net imports appear in the optimization model in Equation 3.20 as S_c. The net import of crops is defined as the nation's imports minus the crops exports of a particular crop. Data from the years 1997-2001 were averaged to obtain typical net import figures for each crop category. The net import data are summarized below in Table 4.9.

Crop	S_c [10^9 kg/yr]
Wheat	2
Rice	-2
Maize	-1
Tubers	4
Total Oils	10
Vegetables	-3

Table 4.9: Average annual net imports of model crop categories into China (Source: FAOSTAT, 2004a).

4.4.4 Growing Degree Days Data

Plants are sensitive to ambient temperature; if it gets too cold a plant will suffer frost or freezing damage and possibly die, too hot and the plant will suffer from heat damage. It is important to know these temperature sensitivities in order to determine when certain crops should be planted.

The original six crop categories considered wheat as one crop category, but many parts of northern China grow winter wheat, which is climatically very different from spring wheat. Therefore the crop category wheat is divided into two separate categories, winter wheat and spring wheat, so that there are a total of seven crop categories considered by the model. However, winter wheat and spring wheat are assumed to have the same expected yield and are counted as one crop for consumption purposes. Winter wheat is unique from the other sequences because it can only be grown over the winter. It is due to this constraint on winter wheat that there are only 42 double-crop sequences and 252 triple-crop sequences possible (simple combinatorics yield 49 double-crop sequences and 343 triple-crop sequences with seven crop categories).

Temperature thresholds for four of the seven crop categories were obtained by combining phenology data (Cui et al., 1984) with the W&M temperature data. Five random grid pixels across China were selected and the corresponding planting and harvesting dates for spring wheat, winter wheat, rice, and maize specified in Cui et al. (1984) were used to determine the average daily

temperature threshold for each crop. This is the temperature at which the plant is able to start growing. Growing Degree Days (GDDs) were summed for the period between the planting and harvesting dates from Cui et al. (1984), using these temperatures for $T_{threshold,c}$ in Equation 3.22 to arrive at the required number of GDDs for each crop, $DD_{req,c}$. The $DD_{req,c}$ values were averaged over the random sampling of pixels to arrive at the final $DD_{req,c}$ values. For winter wheat and rice, the average $T_{threshold}$ values appeared too high using this method, so other sources were used to define $T_{threshold}$ (Miller et al., 2001 and IRRI, 1997).

Temperature thresholds and DD_{req} values for the three remaining crop categories, vegetables, tubers and oil crops, were obtained through an extensive literature review. Since the sub-species of crops within each of these categories have very different properties, the DD_{req} and $T_{threshold}$ for each crop can be very different. An estimate for each model crop category was made based on the available data. The sources for vegetables are Bauder (1999), Miller et al. (2001), Angima (2003), Stilwell and Portas (1978), Scully and Waines (1988), Miller et al. (2002), and Delahaut and Newenhouse (2004). Growing requirement data for tubers is available from Bauder (1999), University of Nebraska-Lincoln (2004), and Angima (2003). Finally, data for oil crops was provided by Bauder (1999), University of Nebraska-Lincoln (2004), Miller et al. (2001), and Angima (2003).

Temperature threshold values for each crop, $T_{threshold,c}$, are summarized in Table 4.10 and the values for crop Growing Degree Days requirements, $DD_{req,c}$, are summarized in Table 4.11. The upper threshold on temperature mentioned in Chapter 3 during the discussion of the Growing Degree Days method is assumed a value of 30°C, obtained from Bauder (1999).

Crop	$T_{threshold}$ [°C]
Spring Wheat	4.1
Winter Wheat	0
Rice	13
Maize	12.1
Tuber	7.2
Oil crop	10
Vegetables	10

Table 4.10: Crop temperature thresholds above which growth is possible.

Crop	DD_{req}
Spring Wheat	1223
Winter Wheat	1752
Rice	1039
Maize	1129
Tuber	1833
Oil crop	1450
Vegetables	1100

Table 4.11: Crop Growing Degree Days requirement.

The data presented in this Chapter are used as inputs to the optimization model described in Chapter 3. In the following Chapter we analyze the results of the model and sensitivity to some of the input data. We also explore the implications for irrigation expansion in China. Finally, we investigate reasons why the population results presented are lower than the current Chinese population.

Chapter 5

Model Results

5.1 Performance

The optimization model is composed of three components that perform complex calculations: the Pre-Processor, Flowrouting module, and GAMS model. The Pre-Processor takes about 2 hours to run on a machine with an Intel Pentium IV processor, 1.8 GHz, and 1Gb of RAM. All clock times are reported for this computer system. The Flowrouting module is the most computational intensive part of the model. Since a highly detailed network was needed in order to extend into every pixel of China, the upscaled flow direction component of the VB executable takes approximately 24 hours to run for all of China. The MATLAB code that is run to create the GAMS flowrouting input takes about 30 minutes to complete. Finally, the GAMS model itself takes less than 5 minutes to find a solution. The GAMS model is designed so that it is able to run within the memory limitations of the computer system. The CPLEX solver is typically the most computational and memory efficient of the GAMS linear programming solvers, but a problem this size can easily require more than 1Gb of RAM. The code is included in Appendix A.

5.2 Baseline

5.2.1 Population & Water – Pareto Analysis

The optimization model is run with the constraints of Chapter 3 and the data discussed in Chapter 4 to obtain baseline results. This formulation is designed to

simulate optimal, sustainable agricultural production restricted by current irrigation infrastructure. In order to obtain a unique optimum solution, a Pareto front was established by parametrically varying the sum of all river runoffs. Each total runoff value corresponds to a unique solution because land is allocated in different ways. There exists one maximum population; however, it is possible to obtain this maximum with many runoff values. Eventually, runoff is forced to a point that can no longer feed the maximum population, so the number of people fed begins to decrease. The best use of water is the point with maximum population and highest runoff at that population – the corner point of the Pareto front. This point corresponds to 693,533,510 people fed and 2,443 km^3 of total runoff for all of China. Figure 5-1, the Pareto front, shows the tradeoff between runoff and population fed.

This Pareto front demonstrates peculiar behavior around a population of 200 million. This appears to be the affect of positive net imports of crops into China. The model assumes that current levels of crop imports are held constant, and this point corresponds to where production for some crops goes to zero but positive imports of other crops all for a non-zero population. After this point the model behaves as expected once again. Since some crops have negative net imports (rice, maize, vegetables) it is possible to obtain zero population.

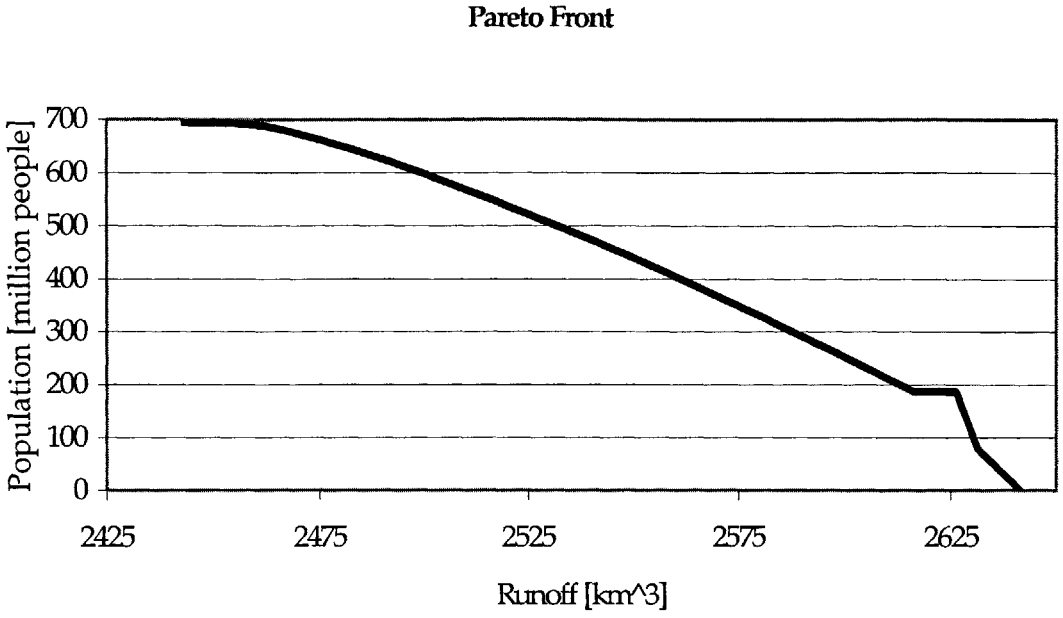


Figure 5-1: Pareto front analysis of tradeoff between total runoff and population fed by modeled agricultural production. The optimum is the point where maximum population is fed with maximum possible runoff to support that population.

At the optimum, evapotranspiration from arable land totaled 485.5 km^3 per yr. This is the total amount of water used for agriculture and it is consistent with water use data from FAO's Aquastat database (2000). For year 2000, Aquastat reports that China used 426.9 km^3 of water for agriculture (Aquastat, 2000).

5.2.2 Crop Production

Crop production results at the optimum are summarized in Table 5.1, below. These production levels are the same as estimated consumption for a population of 693,533,510, according to the crop consumption data of Chapter 4.

Crop	Total Production [10^9 kg]
Wheat	59.74
Rice	72.62
Maize	66.98
Tubers	97.19
Oil crops	27.08
Vegetables	149.53

Table 5.1: Model production results for baseline case with current irrigation capability.

These results tell us nothing about the spatial distribution of the model production results. Pixel crop production results were aggregated to provincial scale and compared with actual crop production data provided by the State Statistical Bureau in the Chinese Statistical Yearbook (1998-2003). The Yearbook production data is averaged over the years 1998-2003. It should be noted that some provinces may not accurately report their data, so any comparison must take this into account. The actual production data and model production results for each of the six main crop categories are shown in Figures 5-3 through 5-7. The Yearbook data does not contain information for the vegetable crop categories, so no comparison can be made in Figure 5-7. Figure 5-2 shows a provincial map of China, for reference.

Figure 5-2: Provincial map of China.

Wheat production is shown in Figure 5-3. The model underestimates production in the northern provinces of Neimenggu (Inner Mongolia), Gansu, Shanxi, Shaanxi, and Hebei. Modeled production in the western autonomous region, Xinjiang, is dramatically lower than the Yearbook production data report.

However, the model appears to overestimate production in the southern provinces of Guizhou, Guangxi, Hunan, Jiangxi, Zhejiang, and Fujian. These discrepancies may be due to the fact that the south can grow two or three crops of wheat in a year, or perhaps in combination with another crop, because of the warmer climate. Also, production in the north may be higher than the model's sustainable prediction because groundwater is being used unsustainable (mined).

Figure 5-4 depicts the actual and modeled production of rice in China. The model appears to do a better job simulating rice production than wheat production. However, the model severely underestimates production from the northeastern provinces of Heilongjiang, Jilin, and Liaoning. These provinces are in the semi-arid northeast where water availability and temperature limitations may be affecting the model's ability to grow rice. Keep in mind that rice has a much higher temperature requirement than wheat and usually requires a lot more water, so it is unable to grow in areas where wheat might be able.

Maize production from the Yearbook data and the model results are compared in Figure 5-5. As with wheat and rice, the model appears to underestimate production in the north and northeastern provinces, most noticeably in Neimenggu, Heilongjiang, Jilin, and Hebei provinces. The model also overestimates production in the southeastern provinces of Jiangxi, Zhejiang, and Fujian. This could be due to the high temperature threshold of maize, which makes the southern areas a more favorable climate for growth in the model despite actual practice in China.

The model most accurately reproduces tuber production, as shown in Figure 5-6. Although production in the northeaster provinces is underestimated, the model predicts the heaviest production in Sichuan province accurately and matches production in other sourthern provinces adequately. Production in Guangxi and Jiangxi provinces appears to be overestimated, but this may be the results of missing data from the Yearbook data.

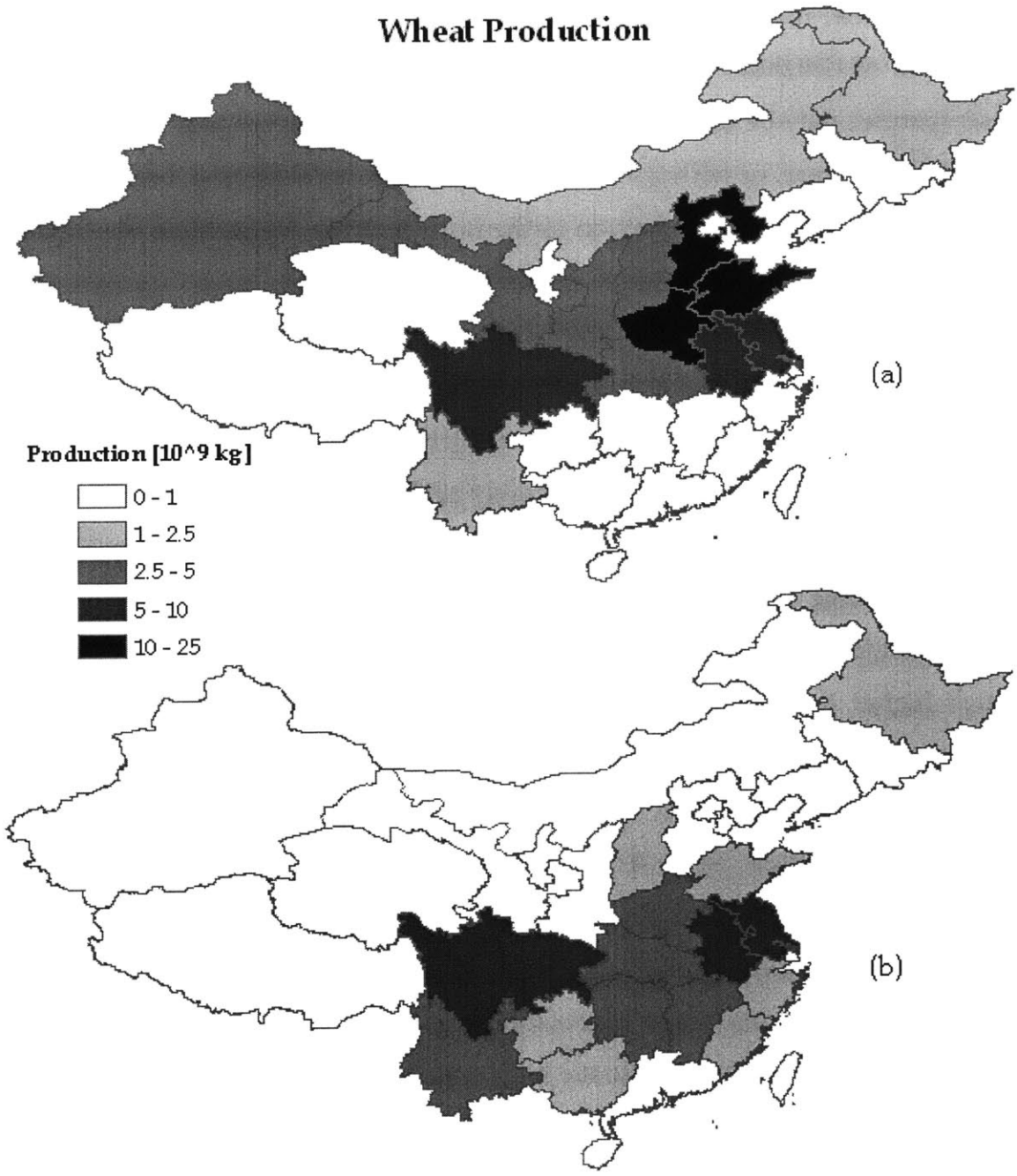


Figure 5-3: Wheat production (a) actual, (b) model results.

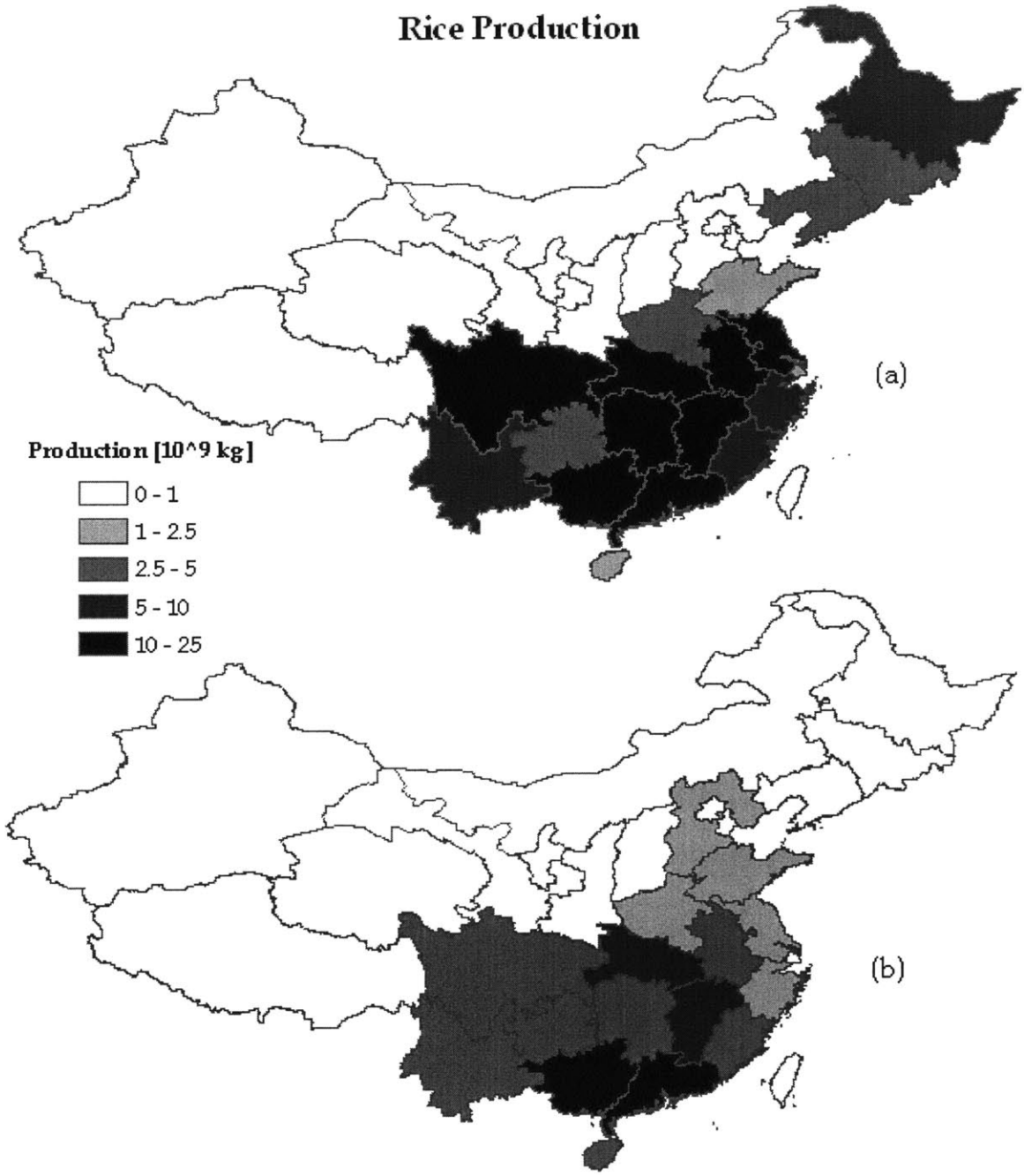


Figure 5-4: Rice production (a) actual, (b) modeled results.

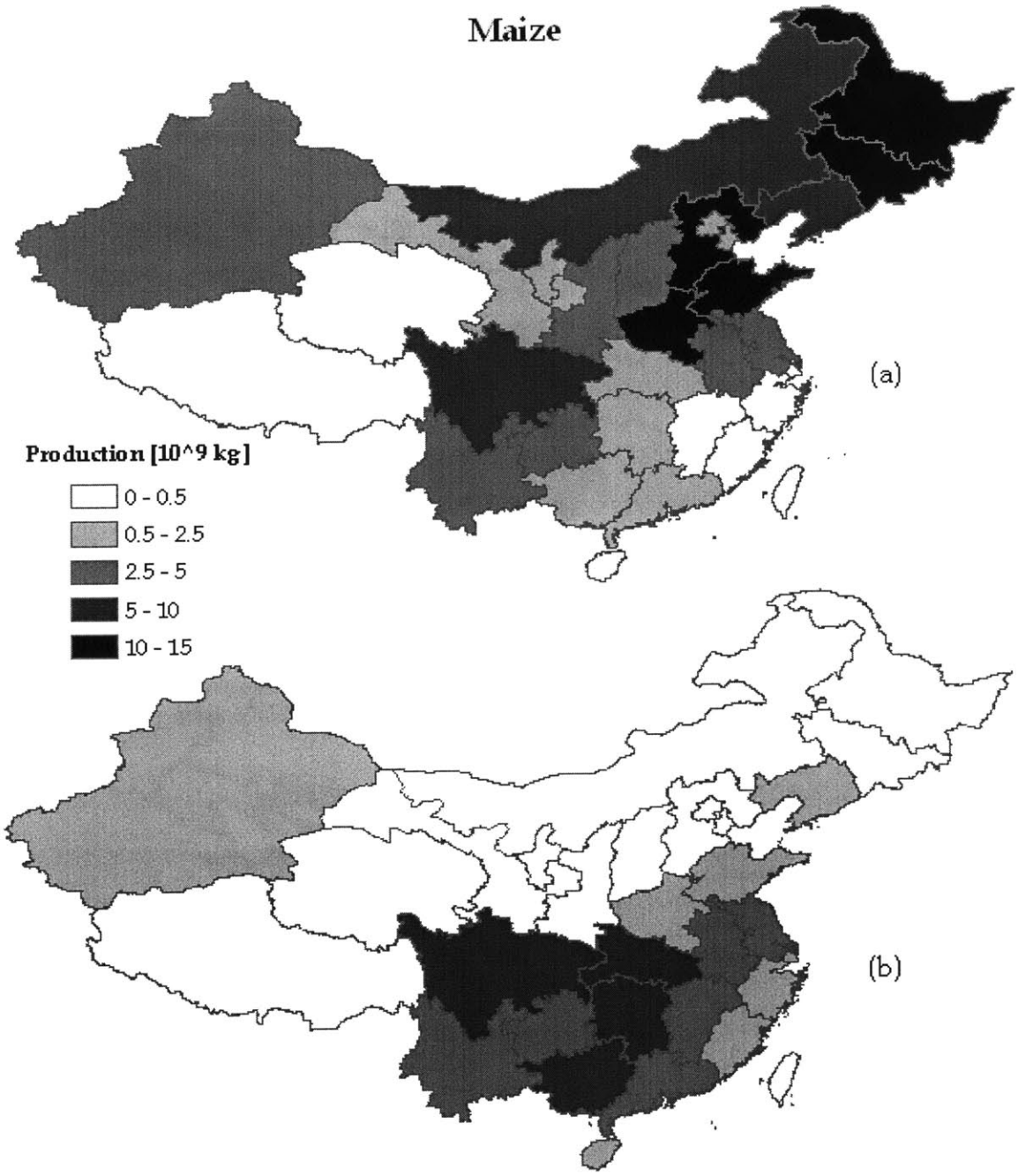


Figure 5-5: Maize production (a) actual, (b) model results.

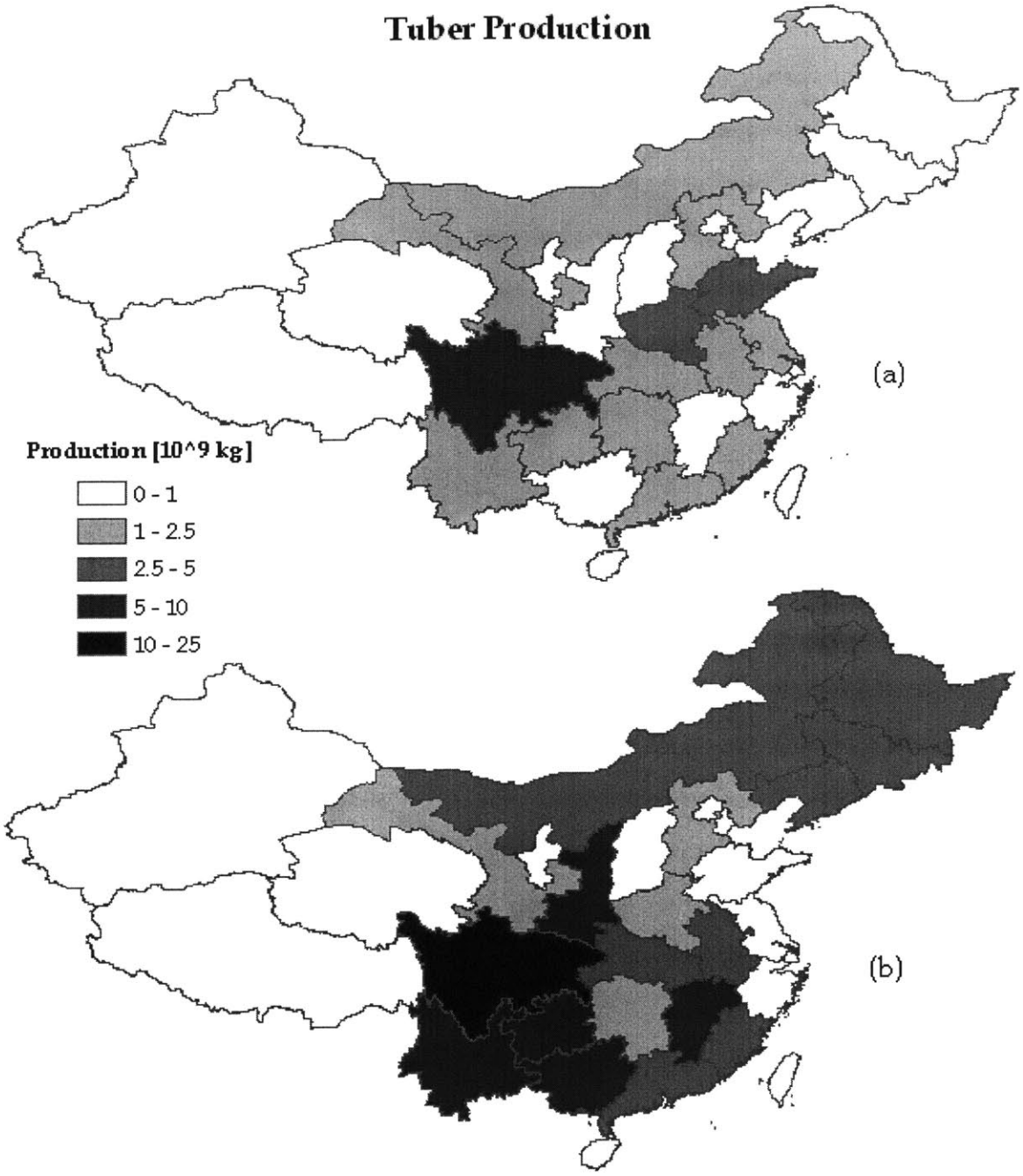


Figure 5-6: Tuber production (a) actual, (b) model results.

Oil production is shown in Figure 5-7. Once again, the model underestimates production in the northern provinces Neimenggu and Gansu. It models oil production in the southern half of the country reasonably well, with heavy production in Sichuan as is seen in the actual data.

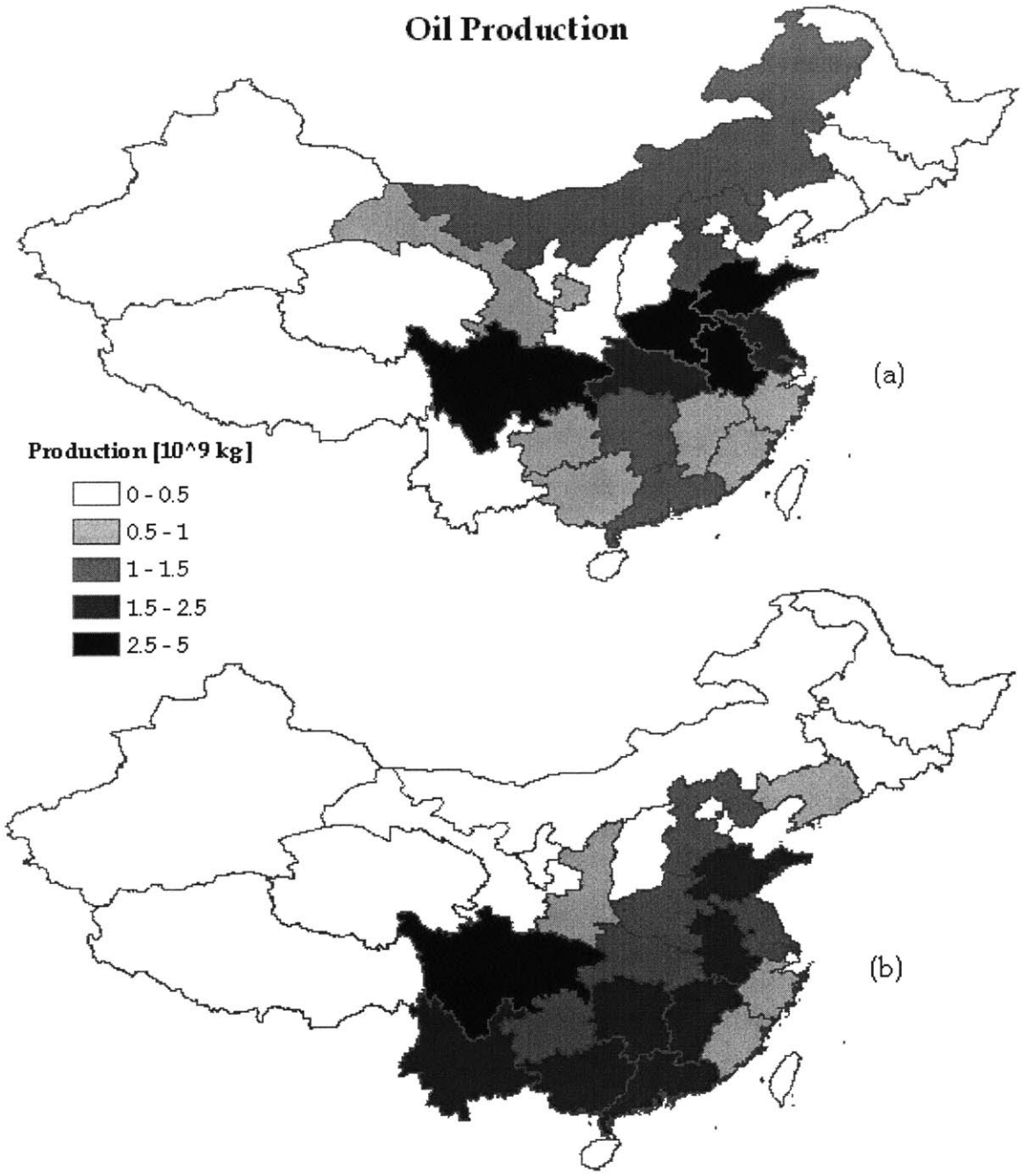


Figure 5-7: Oil crop production (a) actual, (b) model results.

Finally, Figure 5-8 shows modeled vegetable production. An evaluation of vegetable production cannot be made since the Yearbook agricultural data does not contain data for the crops considered in this model crop category. It should be noted that the heaviest production occurs in the southern provinces, most

notably within the Yangtze River basin. Very little vegetable production occurs in the north and western portions of the country.

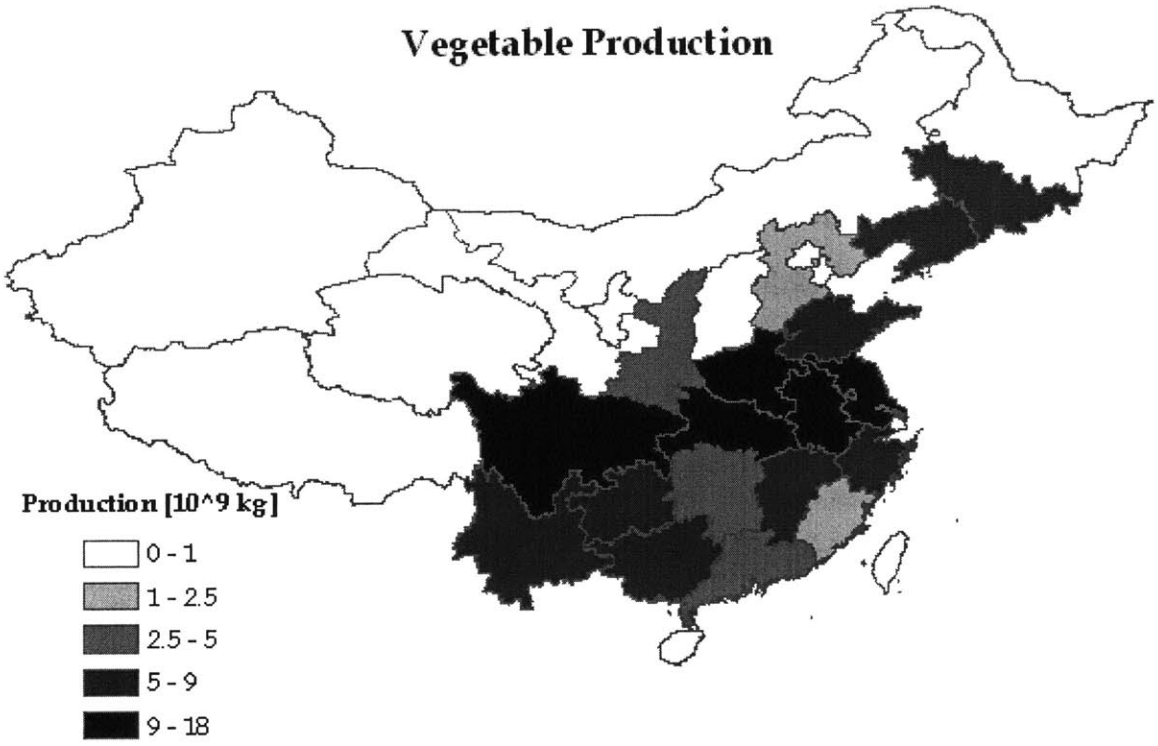


Figure 5-8: Vegetable production model results.

5.2.3 Optimal Irrigation Scenario

It is possible that the level of irrigation infrastructure within China will be expanded in the future, so we explored this possibility by removing the constraint on irrigated land represented in Equation 3.17. The upper bound for future irrigated land is, instead, the total cropland. The model is run with all the same data inputs as for the previous case. With this new assumption about irrigated land, the model is able to feed 828,200,000 people and uses 674 km^3 of water with this scenario. This is an increase of 134,666,490 people fed and 188.5 km^3 of water used for agriculture. These results suggest that expansion of the Chinese irrigation infrastructure could be used to increase agricultural production in order to feed more people.

5.3 Sensitivity Analysis

It was discovered while performing numerous runs of the optimization model that it is highly sensitive to several data inputs, in particular the crop consumption factors and yield factors (J. M. Herrmann, personal communication, April 26, 2004). As a result of this behavior, a simple set of sensitivity tests were performed to better understand the nature of the model's sensitivity to crop consumption factors, crop yields, and the Growing Degree Days requirements.

5.3.1 Crop Consumption

Crop consumption values are used in the model to define the per capita use of crops in China. It is hard to measure the uncertainty in the crop consumption values since they are reported by the FAO with no quantification of error or uncertainty. Sensitivity is conducted by varying each factor individually while keeping all other input values the same. Consumption factors were varied at 50, 75, 125, and 150% of their baseline values presented in Chapter 4. The model showed no sensitivity to wheat or vegetable consumption, but was sensitive to rice, maize, tuber, and oil consumption factors.

The lack of sensitivity to wheat is most likely due to the relatively low water demand of wheat. The model is not sensitive to vegetable crops because they have relatively high yield values. This allows vegetables to be more rapidly produced than the other crop categories.

The model responds nearly linearly to rice, maize and oil consumption, but nonlinearly to tuber consumption. These sensitivities are shown in Figures 5-9 (rice, maize, and tubers) and 5-10 (oil). Decreased tuber consumption values do not increase the supported population, as would be expected, but increased tuber consumption decreases the population fed. This may indicate that other crops, such as oil crops, are limiting the potential for a larger supported population. The scale of Figure 5-10 is larger than Figure 5-9, demonstrating the increased sensitivity of the model to the oil consumption factor.

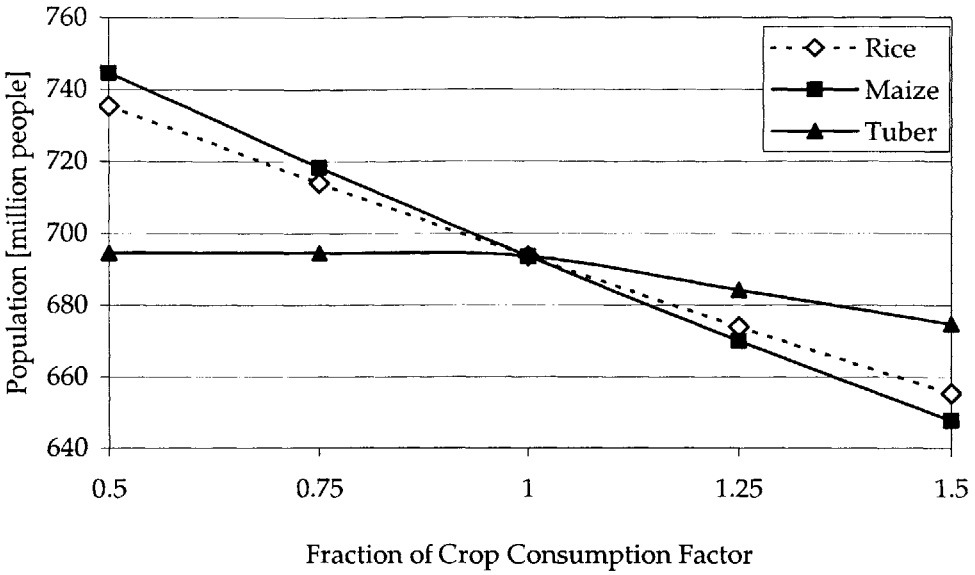


Figure 5-9: Sensitivity analysis of rice, maize, and tuber consumption factors.

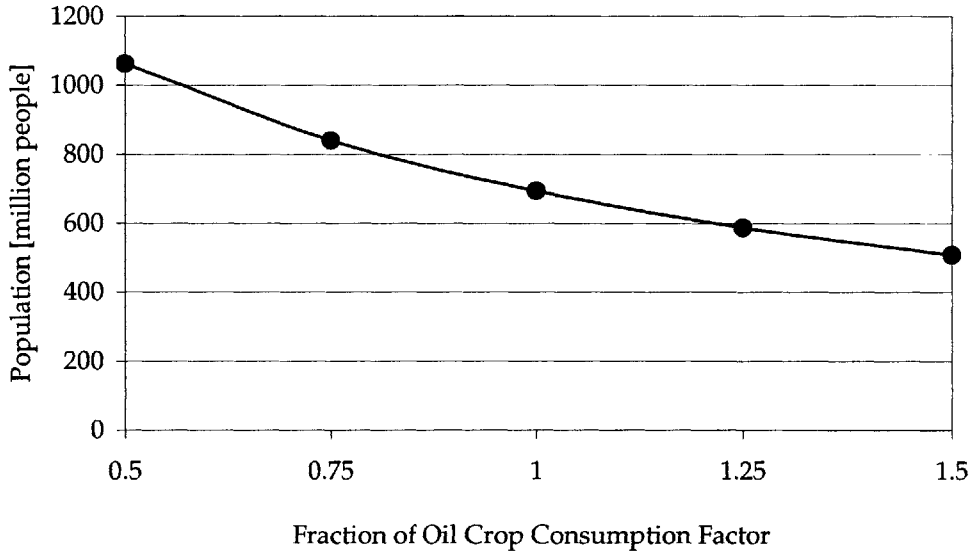


Figure 5-10: Sensitivity analysis of oil crop consumption factor.

5.3.2 Crop Yield

Since production is linearly proportional to the crop yield values [mass/area], it is logical that the model is highly sensitive to the crop yield values. Similar to the crop consumption factors, it is hard to quantify the uncertainty in the yield values. Instead of trying to understand the affect that uncertainty may play in

the model, we were concerned with the potential for improvements in agricultural production in the future. Using the United States as a role model for a highly efficient agricultural country, we compared FAOSTAT crop yields from China and the U.S. and selected the maximum yield values for each crop category. The optimal yield values are summarized in Table 5.2, below.

Crop	Y_c [kg/ha]
Wheat[1]	3.83
Rice[2]	7.14
Maize[2]	8.55
Tubers[2]	39.41
Total Oils[1]	0.53
Vegetables[2]	27.26

Table 5.2: Upper bound crop yield figures, combination of: 1) China and 2) U.S.A. FAOSTAT crop yield figures.

The model was run with these optimal yields and all other inputs at baseline values to obtain an upper bound on the population supported. The model was able to feed 754,910,000 people, only 61,380,000 more people than the baseline case. This scenario used 464.6 km^3 of water for agriculture as compared with 485.5 km^3 in the baseline scenario. This shows that if China were to improve their rice, maize, tuber, and vegetable crop yields to best predictable levels they could feed about 8% more people with less water.

5.3.3 Growing Degree Days

The data for the Growing Degree Days requirements, $DD_{req,c}$ [GDD/crop season] is susceptible to error and varies widely depending on the particular crop species. Within each crop category there are many possible species that can be grown, and each have different growth requirements. Due to this uncertainty we analyzed the affect of lower $DD_{req,c}$ values on the model results. Lowering the $DD_{req,c}$ values by 25% and keeping all other inputs the same, the model feeds 1.485 billion people and uses 569.8 km^3 of water. This is more than twice the

population fed by the baseline data values and 84 km^3 more water for agriculture. Clearly, the model is highly sensitive to the Growing Degree Days requirement of the crops. If the model $DD_{req,c}$ values are too high it will not feed as many people as is actually possible.

5.4 Questioning Sustainability

The model presented in this paper employs a sustainable approach to feed the largest number of people possible. In Chapter 1 we presented evidence that strongly suggests current water use in China, particularly in the north, is not sustainable. The water table depths are declining which indicates a negative change in storage,

$$\frac{dS_b}{dt} < 0 \tag{5.1}$$

This is inconsistent with the formulation of the optimization model in which the change in storage is zero. Therefore, we cannot expect the results of the model to replicate current agriculture production in China.

Our population results suggest that the maximum population that China can support sustainably is 693 million people. The Chinese population surpassed this point in 1963. Figure 5-11 shows the growth of China's population as a function of time.

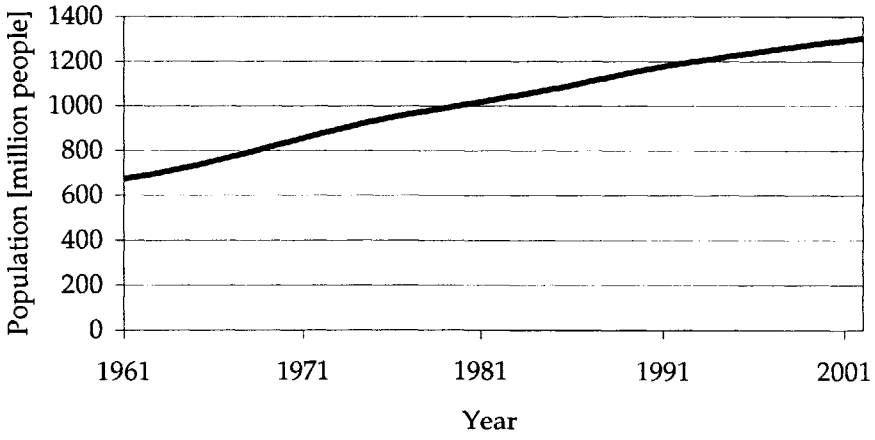


Figure 5-11: China's population growth, 1961-2002. Source: FAOSTAT.

Since declining water tables are an indication of unsustainable use of water, we were interested in knowing when the water tables began to decline in China. The International Water Management Institute (IWMI) compared data of measured and estimated water table depths computed from a simplified groundwater model developed by Kendy et al. (2003). The data is for Luancheng county in Hebei province. These data are shown below in Figure 5-12.

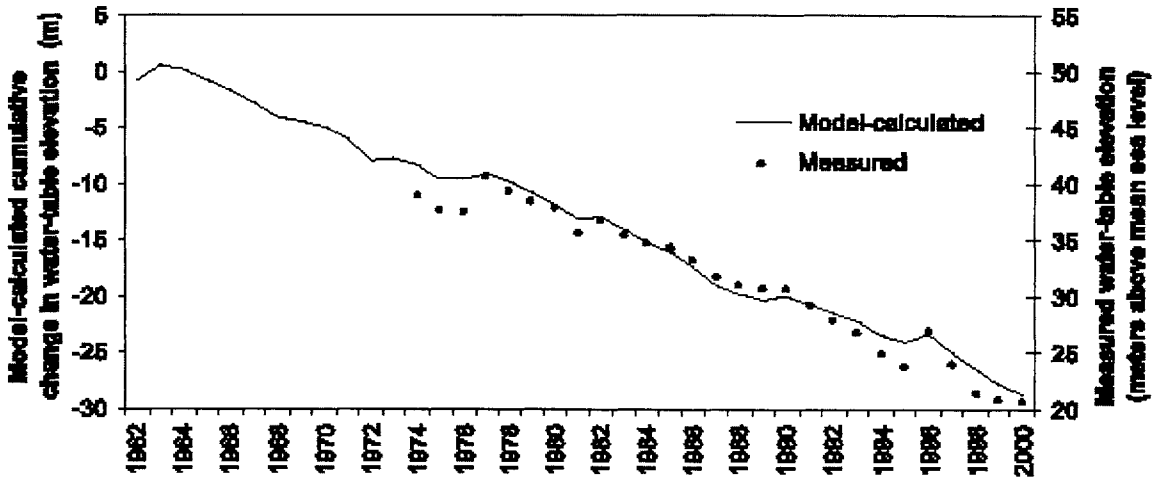


Figure 5-12: Groundwater depletion in Luancheng county, Hebei province, PRC. Source: Kendy et al. (2003).

Interestingly, groundwater depletion in this particular northern county began around the time the Chinese population reached 693 million. Even if our results are overly pessimistic, it is possible that China has been mining groundwater in order to feed its population. It is unlikely that water was being used efficiently in the early 1960's, in terms of supporting their population. It is also unlikely that China is using their water efficiently today. If they were, water intensive crops would be grown primarily in the south as opposed to the north. However, Figures 5-3 to 5-7 show that a large amount of production occurs in the semi-arid north. If our results are correct, this indicates that China has been using water to support an unsustainable population.

There are many assumptions that play a part in the accuracy of this model formulation. In this Chapter we have discussed the sensitivity the model has to some of the data inputs and explored the possibility that sustainable water use is not being achieved in China currently. The following Chapter discusses the limitations inherent to the model and provides suggestions for improvement.

Chapter 6

Conclusions

In this final chapter, we provide a summary of the original contributions of this work as well as suggestions for future research.

6.1 Summary of Original Contributions

The work presented in this thesis makes progress in three main areas. We have: 1) developed an optimization model designed to maximize the population supported by sustainable agriculture, 2) constructed a comprehensive data set for use with the model, and 3) demonstrated how the model is used to determine the most efficient use of resources and shown the model's sensitivity to some of the key inputs.

6.1.1 Optimization Model

We designed an optimization model of sustainable agricultural production in China with the objective of feeding the most people possible. The model provides a better understanding of the limitations of domestic land and water resources on food production in China. It performs a spatially-distributed water balance to determine the most efficient use of land and water. The model makes agronomic decisions as to when crops can be grown and calculates evapotranspiration rates for use in the water balance. The optimization model is composed of two modules, which require GIS, MATLAB, and GAMS software.

The model uses a coarse resolution river network to represent the movement of water within the country. This network accounts for the large-scale effects of

the small-scale structure of river basins and assumes that there is no lateral groundwater flow between basins.

6.1.2 Model Input Data

A comprehensive set of data were compiled and assessed for use with the model, including climatologies of precipitation, evapotranspiration, and temperature. Precipitation and evapotranspiration data are used in the water balance. The model also requires agronomic information to determine when crops are grown and the yield from crops. To estimate the population supported by agriculture production, the model uses consumption data. All sources of data discussed in this thesis are believed to be the best data currently available.

6.1.3 Results and Sensitivity of Model

Preliminary results from the model suggest that China can feed around 700 million people sustainably, given current levels of irrigation. If irrigation infrastructure were expanded to all of the available cropland in the future, China could feed an additional 134.7 million people. This shows that the possibility of feeding more people in the future is realistic. The model results are preliminary and are most likely underestimating China's ability to produce food.

Analysis of changes in several of the model inputs demonstrates the highly sensitive nature of the model. In particular, the model is most influenced by oil crop consumption and the Growing Degree Days requirement. It is possible that further refinement of the oil crop category can alleviate some of this sensitivity.

6.2 Areas for Future Work

6.2.1 Agronomic Concepts

The aggregation of many crops into general model crop categories poses a problem when selecting the agronomic data for these categories. For instance, the Growing Degree Days requirement for crops in the crop category vegetables,

range from 800 to 2000. It is difficult to choose what value within this range is appropriate for the vegetable category. The addition of more crop categories to the model would help resolve this issue, but would in turn increase the computations performed by GAMS. As the model grows it becomes increasingly computational intensive, and could become too large to run on standard computers.

6.2.2 Runoff Data

The model calculates river runoff values and the only constraint on runoff is that it must be greater than zero. River runoff data for China is severely limited. The Global Runoff Data Center (GRDC) distributes stream gauge data for rivers all over the world, but only reports data for about 20 gauges located in China. These gauges are not located at river outlets, which are needed to determine the total runoff to oceans. Runoff data could be compared with the model results to determine if the water balance is producing realistic river flows.

6.2.3 Uncertainty in Precipitation

The three sources of precipitation data, CRU (New et al., 1999), TRMM (DAAC, 2003), and Willmott and Matsuura (2001a) all provide very different estimates of precipitation in China. The Willmott Matsuura precipitation data lie in the middle of the range of reported values and are used in the model. However, if one of the other data sources were used with the model it would certainly produce different results. CRU precipitation would feed less people while TRMM data would feed more people. The uncertainty in the precipitation data is a problem not just for this work but also for all hydrology research. It would be worthwhile to investigate the sensitivity of the model to precipitation and also to resolve the disparity of the three data sets.

6.2.4 Groundwater Issues

The model results may offer an overly pessimistic prediction, but groundwater data suggest that water use in the North China Plain is not currently sustainable. Groundwater mining may not be as severe as our results indicate, but it is reasonable to assume that China is currently using its water resources in an unsustainable manner. Future work should investigate this issue by identifying areas of unsustainable groundwater use and a detailed hydrogeological analysis at regional scales. Additionally, the model can be improved by distinguishing between surface and groundwater basins.

6.2.5 Water Diversion Project

The Ministry of Water Resources in China conducted years of research that suggest China would benefit from the construction of a large scale water diversion project. The project includes several routes in which water would be brought from the water rich south to the water scarce north. With a cost higher than the controversial Three Gorges Dam project, the diversion will be the largest water diversion project in the world, carrying ten-times the amount of water as the next largest diversion (ENN, 2002). The model could be further expanded to allow for water transfers between basins in order to model the diversion. The results of this model would demonstrate the agricultural benefit of the diversion.

6.2.6 Uncertainty of Model Results

The uncertainties of our model results are dependent on the sensitivity of the model to the data and the uncertainty of the data. If the data are uncertain but the model is not sensitive to them, the certainty of the results will not be adversely affected. However, if the model is sensitive to the data and the data are uncertain, then a reasonable amount of uncertainty surrounds the results. This model assumes that all the data are known with certainty, but Monte Carlo analysis can be used to determine the uncertainty of the model results by

estimating a probability distribution of the population. This would give us a quantitative measure of the uncertainty surrounding the model results.

The results presented in this thesis are intended for illustrative purposes more than anything else. They demonstrate how the model can be applied but are not the final results of this project. A quantitative assessment of the uncertainty of the results should be done before any policy implications can be made. In addition, this is a descriptive model not a normative model. The results are intended to tell us something about what is being done currently and give us bounds on how many people China could feed. We do not intend to say what China should do with their land and water, as that is far beyond the scope of this work.

Appendix A

GAMS Optimization Model

```
* GAMS China crop allocation problem
* May, 2004 modified
* Primary decision variable:
*      land(pixel,sequence) -- land (10^6 ha) devoted to crop in pixel and sequence

$offsymxref
$offsymlist
$offlisting
$onempty
option
     limrow=0
     limcol=0          ;
option sysout = off;
option solprint = off;
option LP = Cplex;
option iterlim = 999999 ;
option reslim = 10000 ;

sets
$include china_pixels.txt

alias (pixel,n)

sets
     basin /Amur1, Amur2, Collective, DalingHe, FuchunJiang, HaiHe,
     HanJiang, HuangHe, Indus, LancangJiang, LiaoHe, LuanHe, MinJiang, New1,
     NuJiang, SongDa, SongKoi, Tarim, UlungurHe, XiJiang, Yalu, Yangtze,
     YarluZangbuJiang/ ;

sets
$include trib_def.txt
trib(n,pixel);
$include trib_table.txt

set
* pixel set for each basin
$include basinpixel_def.txt
basinpixel(basin,pixel);
$include basinpixel_table.txt

set
```

```
$include sequence.txt
* sequence set for each pixel
$include pixelsequence_dry_def.txt
pixelsequence_dry(pixel,sequence);
$include pixelsequence_dry_table.txt

set
*$include sequence.txt
* sequence set for each pixel
$include pixelsequence_irr_def.txt
pixelsequence_irr(pixel,sequence);
$include pixelsequence_irr_table.txt

parameters
$include et_irr.txt
$include et_dry.txt
$include precip.txt
$include yieldMz.txt
$include yieldWt.txt
$include yieldHw.txt
$include yieldRc.txt
$include yieldVg.txt
$include yieldTu.txt
$include yieldOl.txt
$include china_area.txt
$include china_croparea.txt
$include irr_area.txt
;

parameter pop_frac(basin)      fraction of nation's population in basin
      /Amur1         0.04845387
      Amur2          0.002922136
      Collective     0.301271456
      DalingHe       0.007119527
      FuchunJiang    0.015607333
      HaiHe          0.050254027
      HanJiang       0.007971471
      HuangHe        0.092314905
      Indus          0.000325252
      LancangJiang   0.006048997
      LiaoHe         0.014749192
      LuanHe         0.013345621
      MinJiang       0.012670822
      New1           0.019826816
      NuJiang        0.003584875
      SongDa         0.001291564
      SongKoi        0.003657275
      Tarim          0.014990045
      UlungurHe      0.000576816
      XiJiang        0.060941709
      Yalu           0.004645787
      Yangtze        0.316998072
```

```
    YarluZangbuJiang          0.00043243 /
;

scalars
* Updated consumption factors - from FAOSTAT Food Balance, 2000
* see excel spreadsheet, China Per Capita Consumption (4.2.04)
develmaizepc   kg per person per yr /95.14/
develwheatpc   kg per person per yr /89.02/
develricepc    kg per person per yr  /101.83/
develvegepc    kg per person per yr  /211.28/
develtubepc    kg per person per yr  /145.90/
develoilpc     kg per person per yr   /53.46/

* M&I consumption in 2001 is 2.35E10 m^3, and population is 1.56E08.
develmipc      developed M&I consumption (m^3 per person per year) /129/
* irrigation efficiency
effc   /0.40/
flow_frac  fraction of basin precipitation in runoff /0/
* conversion from (10^6 ha)(mm yr^-1)
* to (10^6 m^3 yr^-1)
conv1 conversion factor 1      /10/
* conversion from (10^9 kg yr^-1) to (10^6 people)(kg person^-1 yr^-1)
conv2 conversion factor 2      /1000/
* conversion from (sq kilometers) to (10^6 ha)
conv3 conversion factor 3     /0.0001/
* conversion from (10^6 m^3 per yr) to (km^3 per yr)
conv4 conversion factor 4       /0.001/
* based on the sum of a previous run's output
Ball 10^6 m^3 per yr   /0/
**** Import Data ****
import_rice     /-2/
import_wheat     /2/
import_maize     /-1/
import_vege      /-3/
import_tube      /4/
import_oil      /10/
;
positive variables
land_irr(pixel,sequence)     irrigated land use in pixel for sequence (10^6 ha)
land_dry(pixel,sequence)     land use in pixel for sequence (10^6 ha)
etpixel_nnn(pixel)       ET in each pixel for sequence nnn (10^6 m^3 per yr)
etpixel(pixel)          ET in each pixel for all sequences except nnn (10^6 m^3 per yr)
runoff(basin)           runoff water of each basin (10^6 m^3 per year)
mandi(basin)            m and i water of each basin (10^6 m^3 per yr)
tot_maize              total maize yield (10^9 kg per yr)
tot_wheat              total wheat yield (10^9 kg per yr)
tot_rice              total rice yield (10^9 kg per yr)
tot_vege               total vege yield (10^9 kg per yr)
tot_tube               total tuber yield (10^9 kg per yr)
tot_oil               total oil yield (10^9 kg per yr)
ET_tot_arable              total volume of water evapotranspired by arable land (km^3 per yr)
ET_tot_nnn                 total volume of water evapotranspired by non-arable land (km^3 per yr)
```

```
Precip_tot              total precipitation
runoff_sum              sum of all basin runoffs [10^6 m^3 per yr]
yieldmaize(pixel)       total maize yield from each pixel (10^9 kg per yr)
yieldwheat(pixel)       total wheat yield from each pixel (10^9 kg per yr)
yieldw_wt(pixel)        total winter wheat yield from each pixel (10^9 kg per yr)
yieldrice(pixel)        total rice yield from each pixel (10^9 kg per yr)
yieldvege(pixel)        total vege yield from each pixel (10^9 kg per yr)
yieldtube(pixel)        total tuber yield from each pixel (10^9 kg per yr)
yieldoil(pixel)         total oil yield from each pixel (10^9 kg per yr)
precipbasin(basin)      Precip over an entire basin (10^6 m^3 per yr)
precipbasin_arable(basin) Arable Precip over an entire basin (10^6 m^3 per yr)
etbasin(basin)          Arable ET over an entire basin (10^6 m^3 per yr)
etbasin_nnn(basin)      Non-arable ET over an entire basin (10^6 m^3 per yr)
rice_print(pixel)       total wheat production (kg per yr)
maize_print(pixel)      total wheat production (kg per yr)
wheat_print(pixel)      total wheat production (kg per yr)
vege_print(pixel)       total wheat production (kg per yr)
tube_print(pixel)       total wheat production (kg per yr)
oil_print(pixel)        total wheat production (kg per yr)
w_wheat_print(pixel)    total winter wheat production (kg per yr)
need_maize              total maize need (10^9 kg per yr)
need_wheat              total wheat need (10^9 kg per yr)
need_rice               total rice need (10^9 kg per yr)
need_vege               total vege need (10^9 kg per yr)
need_tube               total tuber need (10^9 kg per yr)
need_oil                total oil need (10^9 kg per yr)
;

free variables
chinafed                total population fed in all of China (10^6 people)
;

* Initialize all variables
land_dry.l(pixel,'nnn') = area(pixel) ;
etpixel_nnn.l(pixel) = land_dry.l(pixel,'nnn')*et_dry(pixel,'nnn')*conv1 ;
etpixel.l(pixel) = 0 ;
runoff.l(basin) = sum(pixel$basinpixel(basin,pixel), precip(pixel)*area(pixel)*conv3*conv1
      - etpixel.l(pixel)- etpixel_nnn.l(pixel)) ;
mandi.l(basin) = 0;
chinafed.l = 0;
* all land is initialized as non-arable
yieldmaize.l(pixel) = 0 ;
yieldwheat.l(pixel) = 0 ;
yieldw_wt.l(pixel) = 0 ;
yieldrice.l(pixel) = 0 ;
yieldvege.l(pixel) = 0 ;
yieldtube.l(pixel) = 0 ;
yieldoil.l(pixel) = 0 ;


equations
totalland(pixel)        'total land use in each pixel'
```

```
totalarableland(pixel)     'total arable land use in each pixel'
totalirrigableland(pixel)      'total irrigable land use in each pixel'
pixelet_nnn(pixel)        'ET in each pixel for the nnn sequence'
pixelet(pixel)           'ET in each pixel for all sequences except nnn'
maizeyield(pixel)          'maize yield in each pixel'
wheatyield(pixel)          'wheat yield in each pixel'
w_wtyield(pixel)           'winter wheat yield in each pixel'
riceyield(pixel)          'rice yield in each pixel'
vegeyield(pixel)           'vegetable yield in each pixel'
tubeyield(pixel)          'tuber yield in each pixel'
oilyield(pixel)           'oil yield in each pixel'
mandiwater(basin)          'm and i water use in each basin'
tribbalance(n)            'water balance for area tributary to pixel n'
basinbalance(basin)        'basin-wide water balance'
basinprecip(basin)         'Volumetric precip in each basin'
basinprecip_arable(basin)  'Volumetric arable precip in each basin'
basinet(basin)            'Volumetric et in each basin'
basinet_nnn(basin)         'Non-arable volumetric et in each basin'
runoff_limit(basin)        'ensure minimum river flow'
maizeused                 'maize used directly and indirectly for food'
wheatused                 'wheat used for food'
riceused                 'rice used for food'
vegeused                  'vegetable used for food'
tubeused                 'tubers used for food'
oilused                  'oil used for food'
people_lim                'Limit population to something feasible'
t_maize                  'total maize yield (10^9 kg per yr)'
t_wheat                  'total wheat yield (10^9 kg per yr)'
t_rice                  'total rice yield (10^9 kg per yr)'
t_vege                   'total vege yield (10^9 kg per yr)'
t_tube                  'total tuber yield (10^9 kg per yr)'
t_oil                   'total oil yield (10^9 kg per yr)'
print_maize(pixel)         'total maize production (kg per yr)'
print_wheat(pixel)         'total wheat production (kg per yr)'
print_rice(pixel)         'total rice production (kg per yr)'
print_vege(pixel)          'total vege production (kg per yr)'
print_tube(pixel)         'total tube production (kg per yr)'
print_oil(pixel)          'total oil production (kg per yr)'
print_w_wheat(pixel)       'total winter wheat production (kg per yr)'
n_maize                  'total maize need (10^9 kg per yr)'
n_wheat                  'total wheat need (10^9 kg per yr)'
n_rice                  'total rice need (10^9 kg per yr)'
n_vege                   'total vege need (10^9 kg per yr)'
n_tube                  'total tuber need (10^9 kg per yr)'
n_oil                   'total oil need (10^9 kg per yr)'
et_volume                 'total volume of ET from arable land (km^3 per yr)'
et_nnn                   'total volume of ET from non arabe land (km^3 per yr)'
tot_precip                'total precipitation'
sum_runoff                 'sum of the basin runoffs [10^6 m^3 per yr]'
waterbenef                 'benef of water'
;
```

```
* enforce the land use limit
totalland(pixel)..          sum(sequence$pixelsequence_irr(pixel,sequence),
land_irr(pixel,sequence))
                    + sum(sequence$pixelsequence_dry(pixel,sequence),
land_dry(pixel,sequence))=e= area(pixel)*conv3 ;

totalarableland(pixel)..      sum(sequence$pixelsequence_irr(pixel,sequence),
land_irr(pixel,sequence))
                    +sum(sequence$pixelsequence_dry(pixel,sequence),
land_dry(pixel,sequence))
                    - land_irr(pixel,'nnn') - land_dry(pixel,'nnn') =l= croparea(pixel)*conv3 ;

totalirrigableland(pixel)..    sum(sequence$pixelsequence_irr(pixel,sequence),
land_irr(pixel,sequence)) - land_irr(pixel,'nnn')
                    =l= irr_area(pixel)*conv3 ;

* compute et in each pixel for 'nnn' sequence
pixelet_nnn(pixel)..          etpixel_nnn(pixel) =e= (land_irr(pixel,'nnn')*et_irr(pixel,'nnn')
+land_dry(pixel,'nnn')*et_dry(pixel,'nnn'))*conv1 ;

* compute et in each pixel from all except the 'nnn' sequences
pixelet(pixel)..            etpixel(pixel) =e= sum(sequence$pixelsequence_irr(pixel,sequence),
land_irr(pixel,sequence)*et_irr(pixel,sequence)*conv1)
                    +sum(sequence$pixelsequence_dry(pixel,sequence),
land_dry(pixel,sequence)*et_dry(pixel,sequence)*conv1)
                    - etpixel_nnn(pixel) ;

* water balance for area tributary to pixel n
* excess can only go downstream
tribbalance(n)..            sum(pixel$trib(n,pixel), precip(pixel)*area(pixel)*conv3*conv1 -
etpixel(pixel) - etpixel_nnn(pixel))+
                    (precip(n)*area(n)*conv3*conv1 - etpixel(n)/effc - etpixel_nnn(n)) =g= 0  ;

* precip - et - runoff must be = zero over basin. This constrains
* amount of land cultivated
basinbalance(basin)..         sum(pixel$basinpixel(basin,pixel),
precip(pixel)*area(pixel)*conv3*conv1
                    - etpixel(pixel)- etpixel_nnn(pixel)) - mandi(basin) - runoff(basin) =e= 0  ;

mandiwater(basin)..          mandi(basin) =e= pop_frac(basin)*chinafed*develmipc  ;

* water per basin
basinprecip(basin)..          precipbasin(basin) =e= sum(pixel$basinpixel(basin,pixel),
precip(pixel)*area(pixel)*conv3*conv1) ;
basinet(basin)..            etbasin(basin) =e= sum(pixel$basinpixel(basin,pixel), etpixel(pixel)) ;
basinet_nnn(basin)..          etbasin_nnn(basin) =e= sum(pixel$basinpixel(basin,pixel),
etpixel_nnn(pixel)) ;
runoff_limit(basin)..         runoff(basin) =g= flow_frac*precipbasin(basin);
basinprecip_arable(basin)..    precipbasin_arable(basin) =e= sum(pixel$basinpixel(basin,pixel),
precip(pixel)*conv1*(area(pixel)*conv3-land_irr(pixel,'nnn')-land_dry(pixel,'nnn'))) ;

* compute yield from each pixel for each crop
```

```
maizeyield(pixel)..          yieldmaize(pixel) =e=
sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Mzyield(sequence))

+sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Mzyield(sequence)
)          ;
wheatyield(pixel)..          yieldwheat(pixel) =e=
sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Wtyield(sequence))

+sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Wtyield(sequence));
w_wtyield(pixel)..           yieldw_wt(pixel) =e=
sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Hwyield(sequence))

+sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Hwyield(sequence));
riceyield(pixel)..           yieldrice(pixel) =e=
sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Rcyield(sequence))

+sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Rcyield(sequence))
;
vegeyield(pixel)..           yieldvege(pixel) =e=
sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Vgyield(sequence))

+sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Vgyield(sequence)
);
tubeyield(pixel)..           yieldtube(pixel) =e=
sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence)*Tuyield(sequence))

+sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)*Tuyield(sequence))
;
oilyield(pixel)..            yieldoil(pixel) =e=
sum(sequence$pixelsequence_irr(pixel,sequence),land_irr(pixel,sequence) *Olyield(sequence))
                    +sum(sequence$pixelsequence_dry(pixel,sequence),land_dry(pixel,sequence)
*Olyield(sequence));

* mass balance for maize consumed
maizeused..  chinafed*develmaizepc =l= (tot_maize + import_maize)*conv2  ;
wheatused..  chinafed*develwheatpc =l= (tot_wheat + import_wheat)*conv2  ;
riceused..   chinafed*develricepc =l= (tot_rice + import_rice)*conv2  ;
vegeused..   chinafed*develvegepc =l= (tot_vege + import_vege)*conv2  ;
tubeused..   chinafed*develtubepc =l= (tot_tube + import_tube)*conv2  ;
oilused..    chinafed*develoilpc =l= (tot_oil + import_oil)*conv2  ;
* mass balance for maize needed
n_maize..  chinafed*develmaizepc =e= (need_maize + import_maize)*conv2  ;
n_wheat..  chinafed*develwheatpc =e= (need_wheat +import_wheat)*conv2  ;
n_rice..   chinafed*develricepc =e= (need_rice +import_rice )*conv2  ;
n_vege..   chinafed*develvegepc =e= (need_vege +import_vege )*conv2  ;
n_tube..   chinafed*develtubepc =e= (need_tube +import_tube )*conv2  ;
n_oil..    chinafed*develoilpc =e= (need_oil + import_oil)*conv2  ;


* Population less than 3 billion
people_lim..    chinafed =l= 3000 ;
```

```
* Total volume of water evapotranspired by arable land [km^3 per yr]
et_nnn..        ET_tot_nnn=e=sum(pixel, etpixel_nnn(pixel) ) ;
et_volume..     ET_tot_arable =e= sum(pixel, etpixel(pixel)) ;
tot_precip..    Precip_tot =e= sum(pixel, precip(pixel)*area(pixel)*conv3*conv1) ;

*****************************************************************
t_maize.. tot_maize =e=  sum(pixel, yieldmaize(pixel)) ;
t_wheat.. tot_wheat =e=  sum(pixel, yieldwheat(pixel) + yieldw_wt(pixel)) ;
t_rice.. tot_rice =e=   sum(pixel, yieldrice(pixel)) ;
t_vege.. tot_vege =e=   sum(pixel, yieldvege(pixel)) ;
t_tube.. tot_tube =e=   sum(pixel, yieldtube(pixel)) ;
t_oil.. tot_oil =e=    sum(pixel, yieldoil(pixel)) ;
*****************************************************************
* units are 100 kg per yr
print_rice(pixel)..   rice_print(pixel) =e= yieldrice(pixel)*10000000 ;
print_maize(pixel)..  maize_print(pixel) =e= yieldmaize(pixel)*10000000 ;
print_wheat(pixel)..  wheat_print(pixel) =e= yieldwheat(pixel)*10000000 ;
print_vege(pixel)..   vege_print(pixel) =e= yieldvege(pixel)*10000000 ;
print_tube(pixel)..   tube_print(pixel) =e= yieldtube(pixel)*10000000 ;
print_oil(pixel)..    oil_print(pixel) =e= yieldoil(pixel)*10000000 ;
print_w_wheat(pixel)..    w_wheat_print(pixel) =e= yieldw_wt(pixel)*10000000 ;

sum_runoff..    runoff_sum =e= sum(basin,runoff(basin)) ;
waterbenef..    runoff_sum =g= Ball ;

model china /all/ ;
china.OptFile = 1;
solve china using lp maximizing chinafed ;

file landout_irr/landout_irr.txt/
put landout_irr;
loop(pixel,
      loop(sequence$pixelsequence_irr(pixel,sequence),
      put @1, pixel.tl, @8, sequence.tl, @14, land_irr.l(pixel,sequence) / )) ;

file landout_dry/landout_dry.txt/
put landout_dry;
loop(pixel,
      loop( sequence$pixelsequence_dry(pixel,sequence),
      put @1, pixel.tl, @8, sequence.tl, @14, land_dry.l(pixel,sequence) / )) ;

file etout/etout.txt/
file basinwaterout/basinwaterout.txt/;
*file scalarout/scalarout.txt/;
file crop_yieldout/crop_yieldout.txt/;
file mzyieldout/mzyieldout.txt/, wtyieldout/wtyieldout.txt/,
   rcyieldout/rcyieldout.txt/, vgyieldout/vgyieldout.txt/,
   tuyieldout/tuyieldout.txt/, olyieldout/olyieldout.txt/, wwyieldout/wwyieldout.txt/;

put etout ;   etout.nd=4;
loop(pixel,
  put @1, pixel.tl, @8, etpixel.l(pixel), @20, etpixel_nnn.l(pixel) / ) ;
```

```
put basinwaterout ;
put 'basin        precip        precip_arable        ET        ET_nnn        m&i        runoff    '/;
put '          (10^6 m^3/yr)   (10^6 m^3/yr)   (10^6 m^3/yr)   (10^6 m^3/yr)   (10^6 m^3/yr)
(10^6 m^3/yr)' /;
loop(basin,
  put @1, basin.tl,  @12, precipbasin.l(basin), @29,precipbasin_arable.l(basin),  @46,
etbasin.l(basin), @63, etbasin_nnn.l(basin), @80, mandi.l(basin),@95, runoff.l(basin)  / ) ;
put /;
put @1, 'Total Fed [10^6 people]', @30, chinafed.l /;
put @1, 'Total Runoff [10^6 m^3 per yr]', runoff_sum.l /;
put @1, 'Total arable ET [km^3]', @30, ET_tot_arable.l /;
put @1, 'Total non-arable ET [km^3]', @30, ET_tot_nnn.l /;
put @1, 'Total precip [km^3]', @30, Precip_tot.l //

put mzyieldout ;   mzyieldout.nd=0;
* put 'yieldmaize(pixel) -- total maize yield from each pixel (100 kg/yr)' /  ;
loop(pixel,
  put @1, pixel.tl, @8, maize_print.l(pixel) / ) ;
put wtyieldout ;   wtyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, wheat_print.l(pixel) / ) ;
put rcyieldout ;   rcyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, rice_print.l(pixel) / ) ;
put vgyieldout ;   vgyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, vege_print.l(pixel) / ) ;
put tuyieldout ;   tuyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, tube_print.l(pixel) / ) ;
put olyieldout ;   olyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, oil_print.l(pixel) / ) ;
put wwyieldout ;   wwyieldout.nd=0;
loop(pixel,
  put @1, pixel.tl, @8, w_wheat_print.l(pixel) / ) ;

put crop_yieldout ;
put '              prod(10^9 kg)    need(10^9 kg)'/;
put @1, 'wheat', @15, tot_wheat.l,@30, need_wheat.l /;
put @1, 'maize', @15, tot_maize.l,@30, need_maize.l /;
put @1, 'rice', @15, tot_rice.l,@30, need_rice.l /;
put @1, 'vegetable', @15, tot_vege.l,@30, need_vege.l /;
put @1, 'tubers', @15, tot_tube.l,@30, need_tube.l /;
put @1, 'oil', @15, tot_oil.l,@30, need_oil.l /;
```

Appendix B

Dietary Component

These tables include summations of actual annual Chinese diet components and the model diet composition.

	1995	1996	1997	1998	1999	2000	2001
Total Calories	2858	2914	2958	2981	2981	2980	2961
Model Calories	2573	2624	2656	2689	2707	2711	2679
% in Model	**90.03**	**90.05**	**89.79**	**90.20**	**90.81**	**90.97**	**90.48**

Table B.1: Daily caloric consumption in actual Chinese diet and from model crop categories.

	1995	1996	1997	1998	1999	2000	2001
Total Protein	75.3	78.2	80.1	81.6	82.7	84.5	85.3
Model Protein	66.5	68.7	70.6	71.5	72.6	74.6	75.2
% in Model	**88.31**	**87.85**	**88.14**	**87.62**	**87.79**	**88.28**	**88.16**

Table B.2: Daily protein consumption in actual Chinese diet and from model crop categories.

	1995	1996	1997	1998	1999	2000	2001
Total Fats	69.5	69.8	73.8	78.2	80.6	82.7	83.8
Modal Fats	68.0	68.1	72	76.4	78.8	80.7	81.9
% in Model	**97.84**	**97.56**	**97.56**	**97.70**	**97.77**	**97.58**	**97.73**

Table B.3: Daily fat consumption in actual Chinese diet and from model crop categories.

Appendix C

Population Distribution

River Basin	Population %
Amur1	4.85
Amur2	0.29
Collective	30.13
DalingHe	0.71
FuchunJiang	1.56
HaiHe	5.03
HanJiang	0.80
HuangHe	9.23
Indus	0.03
LancangJiang	0.60
LiaoHe	1.47
LuanHe	1.33
MinJiang	1.27
New1	1.98
NuJiang	0.36
SongDa	0.13
SongKoi	0.37
Tarim	1.50
UlungurHe	0.06
XiJiang	6.09
Yalu	0.46
Yangtze	31.70
YarluZangbuJiang	0.04

Table C: Percent of national population residing in upscaled river basins of China.

Bibliography

Allen, R.G., L.S. Pereira, D. Raes, and M. Smith. 1998. "Crop evapotranspiration – guidelines for computing crop water requirements." *FAO Irrigation and Drainage Paper 56*, Food and Agriculture Organization of the United Nations, Rome. 23 March 2003. <http://www.fao.org/docrep/X0490E/x0490e00.htm>

Angima, S.D. Mar 2003. "Growing degree days what exactly do they mean?" *Ag Newsline* 1(2):1-2. Available online at < http://outreach.missouri.edu/johnson/Ag%20Newsline/Ag%20Newsline%20Mar%202003.pdf>

Aquastat, 2000. "Water use - China." Last updated 2000. 10 April 2004. Available online at < http://www.fao.org/ag/agl/aglw/aquastat/main/index.stm>

Bauder, J. 1999. "What Controls Wheat Growth? Genetics, Temperature." Agronomy notes no. 225. 5 April 2004. Available online at <http://scarab.msu.montana.edu/agnotes/docs/225.htm>

Brown, L.R. 1995. *Who Will Feed China? Wake-up call for a small planet.* W.W. Norton & Company, New York.

Brown, L.R., and B. Halweil. 1998. *China's Water Shortage Could Shake World Food Security.* World Watch.

Brown, L.R. March 2004. "China's Shrinking Grain Harvest." *Eco-Economy Updates.* Earth Policy Institute. 20 March 2004. Available online <http://www.earth-policy.org/Updates/Update36.htm>

CITAS, 1997. County Boundaries of China. China In Time And Space (CITAS). 10 April 2002. Availabe online at <http://citas.csde.washington.edu/data/socioeconomic.html>

Cui, D. H. Liu, J. Min, J. He. 1984. *Atlas of Phenology for major crops in China.* [In Chinese] Meteorological Press, Beijing, China.

DAAC. 17 April, 2003. *README for TRMM Product 3B43.* Goddard Space Flight Center, Distributed Active Archive Center. 25 April 2003. Available online at <http://daac.gsfc.nasa.gov/hydrology/TRMM_README/TRMM_3B43_readme.shtml>

DAAC. 2004. *MODIS Land Cover Type (MOD 12).* Distributed Active Archive Center (DAAC): Earth Resources Observatory System (EROS) Data Center. 15 March 2004. Available online at <http://modis.gsfc.nasa.gov/data/dataprod/pdf/MOD_12.pdf>

Delahaut, K.A. and A.C. Newenhouse. Growing beans and peas in Wisconsin. A3685. 5 April 2004. Available online at <http://cecommerce.uwex.edu/pdfs/A3685.PDF>

Döll, P., 2004. Digital global map of irrigated areas. Center for Environmental Systems Research, University of Kassel, Germany. 24 June 2002. Available online at <http://www.usf.uni-kassel.de/usf/mitarbeit/homepages/doell/research9.htm>

Döll, P. and S. Siebert. 1999. A digital global map of irrigated areas. Kassel World Water Series 1, Center for Environmental Systems Research, University of Kassel, Germany, 23 pp + Appendix.

Doorenbos, J. and W. Pruitt. 1977. Guideline for predicting crop water requirements, *FAO Irrigation and Drainage Paper 24*, Food and Agriculture Organization of the United Nations, Rome. Revised Version.

Eischeid, J.K., H.F. Diaz, R.S. Bradley and P.D. Jones, cited. 1991. "A comprehensive precipitation data set for global land areas." DOE/ER-6901T-H1, U.S. Department of Energy, Washington DC, 81 pp. [Available from FAO, Viale delle Terme di Carcalla, 00100 Rome, Italy.]

ENN. 27 Nov 2002. "China approves gargantuan project to divert water to arid north." *Environmental News Network.* Associated Press. 12 Dec 2002. <http://www.enn.com/news/wire-stories/2002/11/11272002/ap_49067.asp>.

FAO, cited. 1984. *Agroclimatological Data for Africa.* Food and Agriculture Organization of the United Nations, 723 pp.

FAOSTAT, 2004a. "Nutritional Data – Food Balance Sheets." Last updated June, 2003.

FAOSTAT, 2004b. "Agricultural Data: Agricultural Production – Crops Primary." Last updated February, 2004.

Frolking S, J. Qiu, S. Boles, X. Xiao, J. Liu, Y. Zhuang, C. Li, and X. Qin. 2002. "Combining remote sensing and ground census data to develop new maps of

the distribution of rice agriculture in China." *Global Biogeochemical Cycles*, 16(4), 1091, doi: 10.1029/2001GB001425, 2002.

GAMS, 2004. *The GAMS System*. GAMS Development Corporation. 10 March 2004. Available online at < http://www.gams.com/docs/intro.htm>

Heilig, G. K. 1999. *Can China Feed Itself? A System for Evaluation of Policy Options.* International Institute for Applied Systems Analysis (IIASA). 15 September 2002. Available online at < http://www.iiasa.ac.at/Research/LUC/ ChinaFood/index_s.htm>

Huang, J.K. 1998. "Agricultural policy, development and food security in China." In *Agriculture in China 1949-2030*, ed., T.C. Tso, F. Tuan, and M. Faust. Beltsville, Maryland: IDEALS, Inc. pp. 209-257. As cited in Kendry, E., Molden, D.J., Steenhuis, E.T. and C. Liu. 2003. *Policies Drain the North China Plain: Agricultural Policy and Groundwater Depletion in Luancheng County, 1949-2000.*

IRRI, 1997. Report of the Director General 1997. 5 April 2004. Available online at <http://www.irri.org/science/progsum/pdfs/DGReport97/97dgr_ur&fp.pdf>

Kang, S., B. Gu, T. Du, and J. Zhang. 2003. "Crop coefficient and ratio of transpiration to evapotranspiration of winter wheat and maize in a semi-humid region." *Agricultural Water Management*. 59: 239-254.

Kendy, E., D.J. Molden, T.T. Steenhuis, And C. Liu. 2003. "Policies Drain the North China Plain: Agricultural Policy and Groundwater Depletion in Luancheng County, 1949-2000." *Research Report 71.* International Water Management Institute (IWMI). Colombo, Sri Lanka. 27 February 2004. Available online at <http://www.iwmi.cgiar.org/pubs/pub071/Report71.pdf>

LP DAAC, 2003. *EURASIA LAND COVER CHARACTERISTICS DATA BASE.* U.S. Geological Survey (USGS) & National Aeronautics and Space Administration (NASA) Land Processes (LP) Distributed Active Archive Center (DAAC). 12 April 2002. Available online at < http://edcdaac.usgs.gov/glcc/eadoc2_0.asp>

LP DAAC. 2004. *HYDRO1k Documentation.* Land Processes Data Active Archive Center. 3 June 2003. Available online at <http://edcdaac.usgs.gov/gtopo30/ hydro/readme.asp>

Legates, D. R. and C. J. Willmott, cited. 1990a. "Mean Seasonal and Spatial Variability Global Surface Air Temperature." *Theoretical and Applied Climatology* , 41: 11-21.

Legates, D. R. and C. J. Willmott, cited. 1990b. "Mean Seasonal and Spatial Variability in Gauge-Corrected, Global Precipitation." *International Journal of Climatology*, 10: 111-127.

Li, Y., J. Cui, T. Zhang, and H. Zhao. 2003. "Measurement of evapotranspiration of irrigated spring wheat and maize in a semi-arid region of north China." *Agricultural Water Management*. 61: 1-12.

Miller, P., W. Lanier, and S. Brandt. Jul 2001. *Using Growing Degree Days to Predict Plant Stages*. Montguide. Montana State University Extension Service. 5 April 2004. Available online at < http://www.montana.edu/wwwpb/pubs/mt200103.pdf>

Miller, P.R., B. G. McConkey, G. W. Clayton, S. A. Brandt, J. A. Staricka, A. M. Johnston, G. P. Lafond, B. G. Schatz, D. D. Baltensperger, and K. E. Neill 2002. "Pulse Crop Adaptation in the Northern Great Plains." *Agronomy Journal*. 94(2): 261-272.

Miura, T., A.R. Huete, H. Yoshioka, and B.N. Holben. 2001. An error and sensitivity analysis of atmospheric resistant vegetation indices derived from dark target-based atmospheric correction, *Remote Sens. Environ.*, 78:284-298.

Müller, M.J., cited. 1982. *Selected Climatic Data for a Global Set of Standard Stations for Vegetation Science*. W. Junk, 306 pp.

New, M, M. Hulme, and P. Jones. 1999. "Representing Twentieth-Century Space-Time Climate Variability. Part I: Development of a 1961-1990 Mean Monthly Terrestrial Climatology." *Journal of Climate*. 12: 829-856.

Olivera, F., M. S. Lear, J. S. Famiglietti, and K. Asante, 2002. "Extracting Low-Resoltuion River Networks From High-Resolution Digital Elevation Models." *Water Resources Research*. 38(11):1231

Olivera, F. and R. Raina. 2003. "Development of Large Scale Gridded River Networks from Vector Stream Data." *Journal of the American Water Resources Association (JAWRA)*. 39(5):1235-1248.

Postel, S. 1993, "Water and agriculture." In *Water in Crisis: A Guide to the World's Fresh Water Resources*, ed P.H. Gleick, Oxford University Press, Inc., New York. Chapter 5, p.56-66.

Scully, B., and J.G. Waines. 1988. "Ontogeny and yield response of common and tepary beans to temperature." *Agronomy Journal*. 80(6): 921-925

Shiklomanov, I. A. 1993. "World fresh water resources." In *Water in Crisis: A Guide to the World's Fresh Water Resources*, ed P.H. Gleick, Oxford University Press, Inc., New York. Chapter 2, p.13-24.

Smil, V. 1995. "Who will feed China?" *The China Quarterly* 143:801-813.

State Statistical Bureau of the People's Republic of China. 1990, 1998-2003. *Chinese statistical yearbook*. Hong Kong, Beijing, PRC, and New York.

Statistics Publishing House, cited 1991. China County-Level Data on Population (Census) and Agriculture 1990. In Döll and Siebert, 1999.

Stilwell, M. R. and C. A. Portas. 1978. "Timing in Portuguese Direct Seeded Tomatoes for Industry." *Acta Hort.* (ISHS) 72:201-210

University of Nebraska-Lincoln. 2004. Growing Degree Days and Crop Water Use. Crop Watch. The Institute of Agriculture and Natural Resources Cooperative Extension, University of Nebraska-Lincoln. 5 April 2004. Available online at <http://cropwatch.unl.edu/weather/gdd-et.html>

USAF, cited. 1987. *Station Climatic Summaries: Europe.* United States Air Force, 376 pp.

Willmott, C.J., and K. Matsuura. 1 July 2001a. *Terrestrial Air Temperature and Precipitation: Monthly and Annual Climatologies (Version 3.02)* [online]. Center for Climatic Research, Dept. of Geography, Univ. of Deleware. 13 April 2003. Available online at <http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_clim2.html>.

Willmott, C.J., and K. Matsuura. 1 July 2001b. *Terrestrial Water Balance Data Archive: Regridded Monthly Climatologies (Version 1.02)* [online]. Center for Climatic Research, Dept. of Geography, Univ. of Deleware. 13 April 2003. Available online at <http://climate.geog.udel.edu/~climate/html_pages/README.wb2.html>.

Willmott, C. et al. 1985. "Climatology of the terrestrial seasonal water cycle", *Journal of Climatology* 5: 589-606.

World Bank. 2001. "China: Agenda for Water Sector Strategy for North China, Volume 1, Summary Report." Report 22040-CHA. April 2001. Prepared by Ministry of Water Resources, PRC, the World Bank and AusAID. 27 February 2004. Available on-line at <http://lnweb18.worldbank.org/eap/eap.nsf/Attachments/WaterSectorReport/>.

WMO, cited. 1996. "Climatological normals (CLINO) for the period 1961-1990." WMO/OMM 847, 768 pp. [Available from World Meteorological Organisation, 41, avenue Giuseppe-Motta, CP-2300-1211 Geneva 2, Switzerland.]

WRI, 2003. *Ecosystem Area: Cropland/natural vegetation mosaic / Technical and Source Notes* [online]. World Resources Institute. 31 March 2004 Available online at <http://earthtrends.wri.org/text/AGR/variables/763notes.htm>.

Zhu, Y.H., and X.L. Zheng. 1983. "Shallow groundwater resources of the Huang-Huai-Hai plain." In, *Long-Distance Water Transfer: A Chinese Case Study and International Experiences*, et., A.K. Biswas, D. Zuo, J.E. Nickum and C. Liu. 1 March 2004. Available online at <http://www.unu.edu/unupress/unupbooks/80157e/80157E00.htm>